# Automated image-based tracking and its application in ecology

Anthony I. Dell[1], John A. Bender[2], Kristin Branson[3], Iain D. Couzin[4],
Gonzalo G. de Polavieja[5], Lucas P.J.J. Noldus[6], Alfonso Pérez-Escudero[5],
Pietro Perona[7], Andrew D. Straw[8], Martin Wikelski[9,10], and Ulrich Brose[1]

[1] Systemic Conservation Biology, Department of Biology, Georg-August University Göttingen, Göttingen, Germany
[2] HasOffers Inc., 2220 Western Ave, Seattle, WA, USA
[3] Howard Hughes Medical Institute, Janelia Farm Research Campus, Ashburn, VA, USA
[4] Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA
[5] Instituto Cajal, CSIC, Av. Doctor Arce, 37, Madrid, Spain
[6] Noldus Information Technology BV, Nieuwe Kanaal 5, 6709 PA Wageningen, The Netherlands
[7] Computation and Neural Systems Program, California Institute of Technology, Pasadena, CA, USA
[8] Research Institute of Molecular Pathology (IMP), Vienna, Austria
[9] Max Planck Institute for Ornithology, Radolfzell, Germany
[10] Biology Department, University of Konstanz, Konstanz, Germany

**The behavior of individuals determines the strength and outcome of ecological interactions, which drive population, community, and ecosystem organization. Bio-logging, such as telemetry and animal-borne imaging, provides essential individual viewpoints, tracks, and life histories, but requires capture of individuals and is often impractical to scale. Recent developments in automated image-based tracking offers opportunities to remotely quantify and understand individual behavior at scales and resolutions not previously possible, providing an essential supplement to other tracking methodologies in ecology. Automated image-based tracking should continue to advance the field of ecology by enabling better understanding of the linkages between individual and higher-level ecological processes, via high-throughput quantitative analysis of complex ecological patterns and processes across scales, including analysis of environmental drivers.**

## Measuring behavior

Individual behavior (see Glossary) underlies almost all aspects of ecology [1–5]. Accurate and highly resolved behavioral data are therefore critical for obtaining a mechanistic and predictive understanding of ecological systems [5]. Historically, direct observation by trained biologists was used to quantify behavior [6,7]. However, the extent and resolution to which direct observations can be made is highly constrained [8] and the number of individuals that can be observed simultaneously is small. In addition, an exact record of events is not preserved, only the biologist's subjective account of them.

Recent technological advances in tracking now make it possible to collect large amounts of highly precise and accurate behavioral data. For many organisms equipment can be attached that provide information about the

### Glossary

**Background subtraction:** a method used by software to compare the current video frame with a stored picture of the background; any pixel of the current frame that is significantly different from the corresponding pixel in the background is likely to be associated with the body of an animal. Useful in situations where the background is unchanging, for example, when the surface of the background is rigid and lighting does not change.

**Behavior:** the actions of individuals, often in response to stimuli. Behavior can involve movement of the individual's body through space, such as walking or chasing, or can occur while the animal is stationary, such as grooming or eating.

**Bio-logging:** attachment or implantation of equipment to organisms to provide information about their identity, location, behavior, or physiology (e.g., global positioning systems, accelerometers, video cameras, telemetry tags).

**Ecological interaction:** any interaction between an organism and its environment, or between two organisms (i.e., including interactions between conspecifics).

**Fingerprinting:** a method used to identify unmarked individuals using natural variation in their physical and/or behavioral appearance. The method works by transforming the images of each individual into a characteristic 'fingerprint', which can then be used to distinguish individual organisms both within and across videos.

**FPS (frames per second):** the number of frames in an image sequence collected per second.

**Image:** any measurement of the spatiotemporal position or pose of organisms that can be recast into a digital image and analyzed using computer vision techniques (see Box 2).

**Machine learning:** a set of techniques that allow computer software to learn from empirical data, user assumptions, or manual annotation. These approaches are becoming increasingly common in the analysis of behavior, where users can tag behavior in short sequences of images and the software can predict occurrences of these behaviors throughout the entire image sequence.

**Marking:** the attachment of artificial 'marks' to organisms to maintain their identity, such as paint or barcodes.

**Occlusion:** when the view of any individual in an image is disrupted either by another individual or physical habitat (i.e., the occluding object lies in a straight line between the focus individual and the camera).

**Pixel:** a physical point in a 2D digital image, and therefore the smallest controllable element of a picture represented on the screen. The equivalent of a pixel in 3D space is a voxel.

**Pose:** any additional geometrical quantity of interest other than the center of the main body of the animal, such as orientation, cell positions, body curvature, etc.

**Position:** the center of body mass of an individual in time and space.

**Resolution:** the number of pixels/voxels in a digital image.

---

**Box 1. Ecological insights from automated image-based tracking**

We see three key areas where considerable intellectual progress has been made in ecology using automated image-based tracking. First, the kinematics of animal behavior [17–19,23,24,34,42,57,66,69,70,74, 76–81], including the role of the internal state of animals, such as their physiology or genes, and the external environment. Recent breakthroughs in remote quantification of physical landscapes [58–60] and 3D imaging [29] should be especially helpful for these questions. Second, collective behavior in animal groups [1,26,33,38,40,43,45,62, 82,83], including understanding how information about the physical and biological environment transfers between individuals. Generally, this research centers on intraspecific groups comprising large numbers of similar sized individuals. Third, determinants of social behavior [8,27–29,31,53,54,67,71,73,84]. Research in this last category usually focuses on a small number of individuals, because identifying the detailed pose required for automated behavioral analysis is difficult in larger groups. Tracking over short durations (minutes) has aided in our understanding of the genetic basis of social behavior, such as aggression or courtship [8,85], where the high throughput that automation allows provides enhanced power for uncovering patterns in behavioral data [27]. Research over longer times can uncover complex temporal linkages between social behaviors [8,28], and experiments over the order of weeks provide unique insight into

the social and behavioral development of individuals in intraspecific groups [31,53,54].

Enormous potential exists for automated image-based tracking to address other key issues in ecology. One area we expect significant growth is in the study of interspecific interactions, which are critical to ecological systems [1–5]. For example, biologists recently used automated analysis of sonar images to reveal how coordinated hunting by predators leads to increased fragmentation and irregularities in the spatial structure of prey groups, and thus inhibition of information transfer among prey [4]. Laboratory research alone provides much scope for experimentally testing basic ideas about ecology, such as the role of body size or predator density in determining trophic interaction strength (Movie S3 in the supplementary material online) (A.I. Dell, unpublished). Image-based tracking can also address more applied questions, such as the role of fragmentation in population dynamics (A.I. Dell, unpublished) or determining the size of animal populations that are historically difficult to measure [52]. Integrating automated tracking techniques into images already collected by trigger-based cameras to assess species occurrence and population abundances [21] would provide important information about the behavior of organisms in natural ecosystems.

---

individuals' spatiotemporal position, orientation, and physiology. This 'bio-logging' allows remote reconstruction of behavior over large spatiotemporal extents, providing essential individual viewpoints, tracks, and life histories, and thus important ecological and evolutionary insights [9–11]. Image-based tracking, for example with video, is another tracking method that shows great potential in ecology. Similar to bio-logging, image-based tracking involves digital recording of data, meaning an objective view of events is maintained, increasing repeatability of studies, and allowing biologists to mine data for quantities not originally considered. Image-based tracking can be used when individuals are too small to attach bio-loggers, or if the equipment itself changes behavior, and all visible and sufficiently resolved individuals within the imaged area can be tracked, not just those with loggers attached. Also, image-based tracking generally allows for higher spatiotemporal resolution of behavioral data than bio-logging, and many imaging methods allow extraction of quantitative information about the environment, such as its temperature or topography. Currently, constraints on the acquisition, processing, and storage of digital information limit the spatiotemporal extent of image-based tracking, and extracting the position and pose of every individual in each image is difficult in complex habitat and at high densities. Nonetheless, constraints are rapidly being overcome and image-based tracking now provides a valuable tool to undertake rigorous hypothesis-driven research in ecology (Box 1). Here we review the state-of-the-art of image-based tracking, its strengths and limitations when applied to ecological research, and its application to solve relevant ecological questions.

**Automated image-based tracking**

Initial applications of image-based tracking required manual analysis [12,13], which is effort intensive, often leads to poor spatiotemporal resolution, and is open to observer effects such as subjective decisions about which information to record. Recent advances in automation are overcoming these issues [14–16], and there now exist several

image-based systems capable of extracting individual behavior with minimal or zero manual intervention (Table S1 in the supplementary material online). Tracking over ecologically relevant spatiotemporal scales is becoming easier, owing to advances in imaging and computing technologies, and by the development of software that can track in real time [17–19] and recognize individuals across image sequences [20,21]. Biologists now employ a wide range of imaging methods (e.g., near infrared, thermal infrared, sonar, 3D) that permit tracking in environments where optical video is unsuitable (Box 2). To date, automated image-based tracking has primarily been undertaken in the laboratory, where biologists have examined genetic and physiochemical drivers of behavior in model species (Table S1 in the supplementary material online) (Box 1). However, the past decade has seen expansion of these methods into the field, and automated image-based tracking has now been undertaken on a wide diversity of species, including plants, worms, spiders, insects, fish, birds, mammals, and more (Table S1 in the supplementary material online).

Automated image-based tracking involves three main steps (Figure 1): (i) acquisition of image sequences (Box 2); (ii) detection of individuals and their pose in each image and appropriate 'linking' of detections in consecutive images to create trajectories through time (Box 3); and (iii) analysis of behavioral data (Box 4). Real-time tracking is performed as images are acquired, removing the need for storing large amounts of digital information [17–19] and allowing researchers to influence the animal's environment in real time through virtual reality, robotics, or other dynamical stimulus regimes [22–24]. Even under controlled laboratory conditions with small numbers of individuals, automated image-based tracking is a difficult computer vision problem. Biological organisms are highly deformable objects which behave in unconstrained and variable ways [25], and the environments within which they exist are complex and dynamic.

Ultimately, in automated image-based tracking there is a trade-off between the difficulty of the tracking problem

## Box 2. Obtaining an image sequence

The first step in automated image-based tracking involves obtaining a machine-readable sequence of images that accurately represents the real world. This translation between the real and digital world is a critical step, and time spent optimizing the image (such as ensuring sufficient contrast between foreground and background) pays substantial dividends during subsequent steps (see Figure 1 in main text). Optical video is commonly used owing to its accessibility and low cost, but other imaging technologies considerably expand the range of environmental contexts within which tracking can be undertaken (Figure I). These include infrared (Figure IA,B), thermal infrared [50] (Figure IC; Movie S7 in the supplementary material online), X-ray microtomography [55] (Figure ID), and sonar [4] (Figure IE; Movie S9 in the supplementary material online). Light-field (Figure IF) and multi-scale gigapixel [86] (Figure IG) imaging should permit tracking and scene reconstruction in 3D from a single image viewpoint. Although frame rates of gigapixel cameras are increasing (S.D. Feller, unpublished), at three frames per minute [86], they are currently too slow for most automated tracking applications. Light-field cameras work at higher frame rates and there are several laboratories exploring if they can be successfully incorporated into automated tracking systems (I.D. Couzin and G.G. de Polavieja, unpublished). Ultimately, decisions about which imaging method to use should be determined by the specific needs of the project.

Automated tracking generally requires a high-contrast image so that computer vision algorithms can adequately discern organisms and their appendages from the surrounding background (Box 3). A common and low-cost method of obtaining such images is to construct an artificial arena for tracking experiments, which is often colored in contrast with the animals, and brightly and uniformly lit with diffuse lighting (Figure IA,B). Deciding on the spatial and temporal resolution of images is also a key consideration. Higher resolutions generally result in better tracking results and more precise quantification of behavior, but bottlenecks during the transmission, storage, and processing of digital information can limit high temporal resolution to low spatial resolution and/or short durations. Constraints on low spatial resolutions can be overcome by integrating output from multiple cameras [18] and should become less important as technology advances. Recording software is another important consideration, such as the choice of codec for encoding and compressing digital data or ensuring that accurate time stamps are obtained and that frames are not silently dropped, and robust open source [87,88] and commercial [Noldus Information Technology, media recorder, 2013 (http://www.noldus.com/media-recorder); Norpix, StreamPix, 2013 (http://www.norpix.com/products/streampix/streampix.php)] options are available.
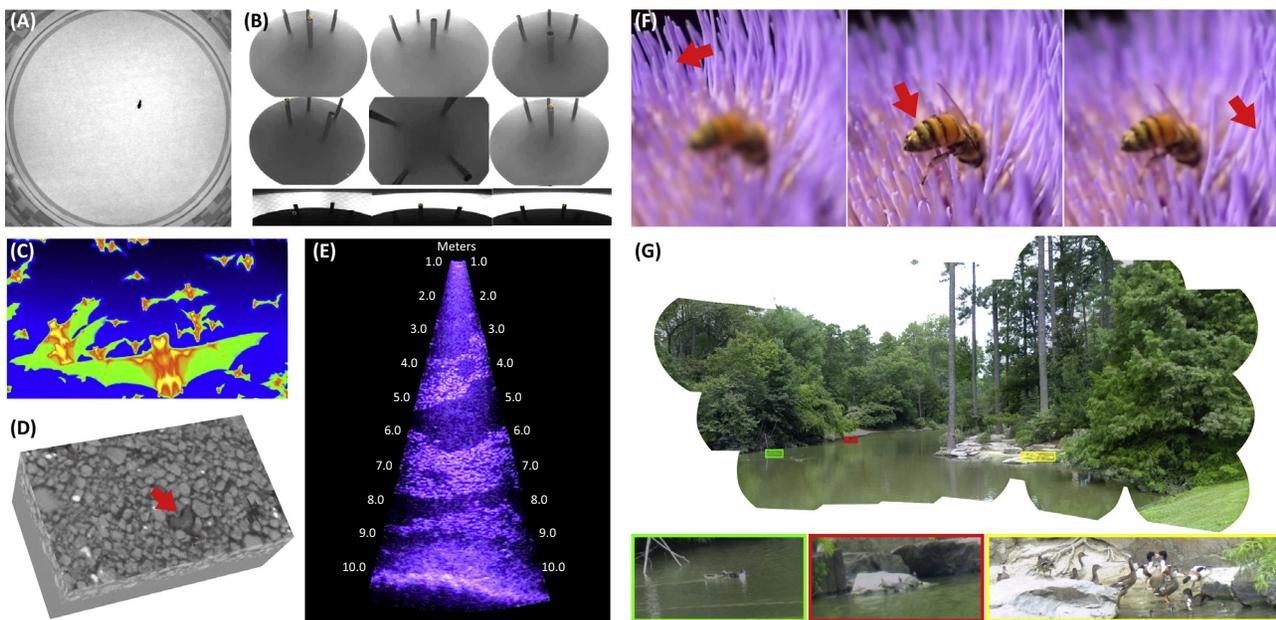


*TRENDS in Ecology & Evolution*

**Figure I**. A growing number of technologies allow capturing of digital images for automated image-based tracking. **(A)** The most common is optical or near infrared video, most often used in simple 2D laboratory settings (left panel in Figure 1) (Movie S1–S4, Movie S5, Movie S10, Movie S11, Movie S14, and Movie S17 in the supplementary material online). **(B)** Images from multiple cameras allow tracking in 3D, even with some degree of habitat complexity present (Movie S6 and Movie S15 in the supplementary material online). **(C)** Thermal imaging allows tracking in complete darkness, but requires that tracked animals have a surface temperature different from the surrounding landscape (Movie S7 in the supplementary material online). **(D)** High-resolution X-ray microtomography permits imaging through complex habitat structure, such as soil (burrowing invertebrate highlighted by red arrow). **(E)** Acoustic imaging (sonar) can also image in habitats where optical video would be unusable, such as this image of predators foraging for schooling bait fish in a turbid estuary [4] (Movie S9 in the supplementary material online). **(F)** Light-field cameras allow for *post-hoc* selection of focal points, thus potentially allowing tracking and construction of the scene in 3D from a single image point. The three panels in (F) were obtained from a single light-field image – each panel representing different focal points (highlighted by a red arrow). **(G)** Newly developed gigapixel technologies also permit capturing of images from a single image point with very high spatial resolutions and at multi-scales, again allowing for 3D tracking from a single image point [86]. The three lower panels in (G) are enlarged sections of the main image. See Acknowledgments for credits and permissions.

(horizontal axis in Figure 2) and the quality of tracking output (vertical axis in Figure 2).

### Difficulty of the tracking problem

Tracking is easiest in laboratory-based systems with a simple environmental landscape and low numbers of individuals (left panel in Figure 2), and most difficult in the field where many individuals from many different species interact across a complex environmental landscape (right panel in Figure 2).

### From individuals to interactions

Monitoring the behavior of individuals as they interact with each other is difficult for several reasons. First,
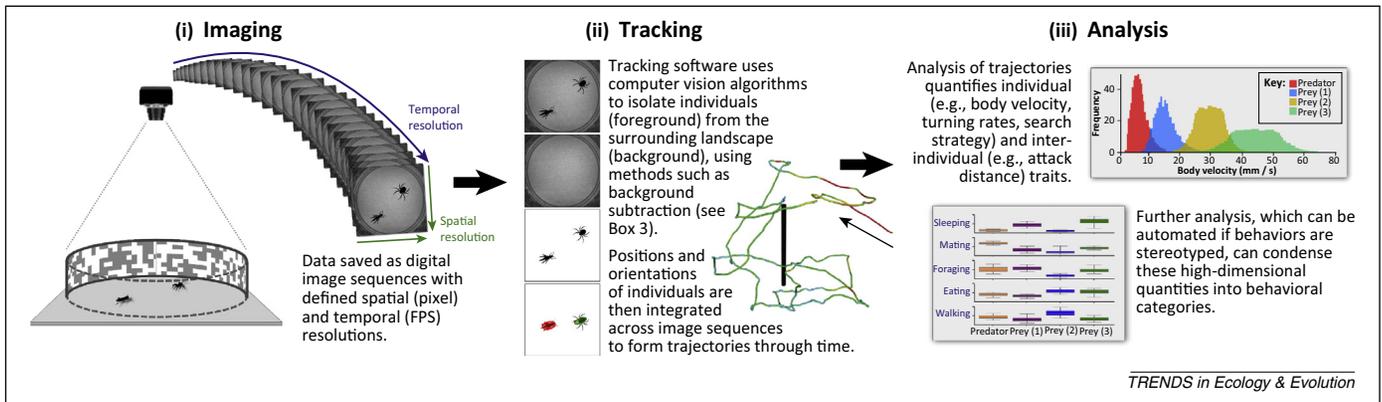
**Figure 1**. The three general steps involved in automated image-based tracking of behavior are: (i) imaging (Box 2); (ii) detection of individuals and their pose in the image and appropriate 'linking' of detections to create separate tracks through time for each individual (Box 3); and (iii) analysis of trajectory and behavioral data (Box 4). To date, imaging is often done in the laboratory (left panel), which can more easily provide a clean, crisp image that minimizes tracking errors. Each of these steps are strongly interlinked and time spent optimizing one step (e.g., imaging) can pay huge dividends in time and effort saved at later steps (e.g., reducing tracking errors).

organisms often move rapidly when interacting (Movie S13 in the supplementary material online), requiring data with high spatiotemporal resolution. Second, because multiple individuals are involved, interactions are prone to occlusions, made especially worse because interactions often involve close physical contact. Occlusions cause identity errors, which are not local but propagate throughout the remaining image sequence. Manual corrections of these errors are labor intensive. Customized automated algorithms which predict identity based on the relative speed and direction of movement can reduce mistakes, and thus dramatically reduce the number of manual interventions needed [26,27], but error propagation is still unavoidable because of the stochastic behavior of organisms [15] (Box 3). 'Fingerprinting' somewhat resolves this problem (see below), but maintaining identities always becomes more difficult as the number of close individuals scales with increasing density. Tracking individuals during occlusions is an additional problem and can be partly overcome when prior knowledge about the shape of the organisms is incorporated into the system [26–28]. Recent approaches utilizing multiple 3D depth cameras are especially useful in this regard [29] (Movie S22 in the supplementary material online) and could eventually be integrated with fingerprinting to assist in resolving identities during occlusions.

Most current attempts to track multiple individuals involve organisms that are similar in size and shape (Table S1 in the supplementary material online). In nature, however, interactions between species often involve individuals that differ greatly in size and shape [30] (Movie S13 in the supplementary material online). Although such differences can be useful for distinguishing individuals [8,20], many tracking systems rely on knowledge about the typical shape of individuals to aid in the segmentation and analysis of images [27,28,31]. Even if shape issues are overcome, it remains a difficult task for computer vision algorithms to separate small animals from the body and appendages of larger animals. Algorithm features allowing tracking of differently sized and shaped organisms, such as more sophisticated contour representations or fingerprinting, would greatly enhance the usefulness of image-based tracking to ecologists (Box 5).

## Tracking in three dimensions

Automated image-based tracking in 2D environments is substantially more straightforward than in 3D (Figure 2). Therefore, many tracking systems are limited to simple 2D arenas and either involve organisms that naturally move in 2D or quasi-2D, or work by constraining normally 3D individuals to only move in 2D. This latter method can be achieved by modifying organisms directly, such as by wing clipping [27], or by using physical boundaries to constrain behavior to near 2D [1,20,27,32,33] (Movie S1, Movie S4, Movie S5, and Movie S10 in the supplementary material online). In nature, however, most organisms incorporate at least some degree of movement in 3D, which influences ecological interactions [3]. Tracking systems designed for 2D can provide some resolution for behavior in a third spatial dimension [34], but ultimately developers must produce tracking systems that can successfully track large numbers of animals in 3D space (Movie S8 in the supplementary material online).

Tracking unconstrained flying or swimming animals can be achieved in several ways, but most often multiple cameras are employed [18,29,35–45] (Movie S6, Movie S8, and Movie S22 in the supplementary material online). Although only two calibrated cameras taking images of the same point in space are required for triangulation, information from additional cameras can incrementally improve localization, especially if some cameras are limited by occlusion or low contrast [18]. Synchronizing multiple cameras requires additional hardware and more complicated software that relates equivalent objects between image sequences; however, this complexity can be hidden from the user by dedicated multi-camera systems [18]. Triangulation is optimized when cameras are positioned with maximally divergent locations, which in the field can introduce problems because arranging unobstructed cameras at multiple locations can be difficult, as can be obtaining multiple views of every location of interest.

Some technologies allow 3D tracking from a single imaging device, which could solve many of these issues. For example, 3D images can be reconstructed from a single image of reflections or shadows on a 3D surface [46,47],

## Box 3. Identifying individuals and behaviors in images

Once a set of suitable images has been obtained (Box 2), the position of individuals, and often their pose, must be automatically computed to form trajectories through time. First, the software must determine whether and where individuals are present in each image. How easily this is done varies with the type and quality of images (Box 2), as well as how accurately each individual's position can be predicted from its previous behavior (see below). Detection is straightforward when the contrast between individuals and the background is substantial, and when the background is known or does not change throughout the entire image sequence is most easily performed by background subtraction (Figure IA–C). The physical complexity of natural systems will ultimately require more advanced techniques, such as those which constantly update their background image [18], or through visual recognition methods [21,63–67], where the distinctive pattern associated with an individual's body and its motion can be recognized against the clutter of the background.

The output of the detection stage is an estimate of the pixels associated with individuals in each image. The position and pose of organisms with stiff and simple-shaped bodies can be computed by fitting a shape contour to the image of the organism [8,27] (Figure ID),

including determining whether clumps of pixels should be separated into multiple individuals (Figure IE–I). The situation is more complex when the body is flexible and multiple degrees of freedom are of interest, such as wing angles or head orientation (Figure IJ). Algorithms for learning and computing an individual's pose is an active area of research, and involves either explicit modeling of the body, or learning associations between image brightness patterns and pose parameters [68,72,76].

Finally, the position of each individual must be linked over multiple frames to form trajectories (Figure IL–P). This is relatively simple for single individuals, although false and missed detections become more likely when detection is problematic. Constructing trajectories for multiple individuals often involves parameterization of a movement model which includes information from previous frames, such as the acceleration of each individual or their preferred direction of motion [89,90]. Movement models also improve the detection phase of tracking, but ultimately suffer from error propagation and thus can be labor intensive. Fingerprinting identifies individuals from their image structure (see main text) and therefore recovers identities after occlusion [20] (Figure IK; Movie S5 in the supplementary material online).
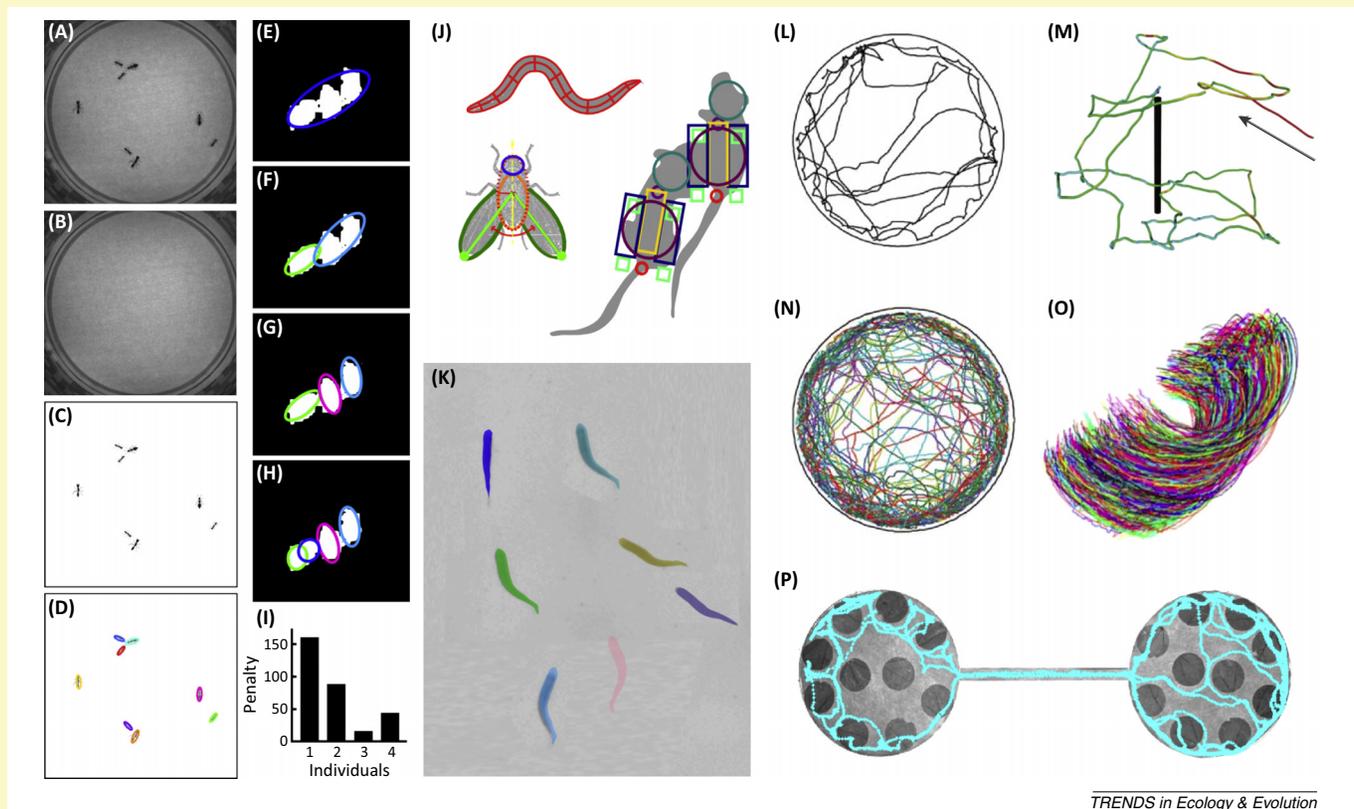


*TRENDS in Ecology & Evolution*

**Figure I**. After imaging (Box 2), computer vision software must automatically detect the position, and sometimes pose, of individuals in the image to create trajectories. **(A–C)** A common approach for detecting individuals is background subtraction, where detection of individuals in raw images is achieved by removing an estimated background-only image, resulting in isolation of foreground pixels. **(D)** Contours, denoting individuals, can then be mapped on to clusters of these foreground pixels. How many individuals are within a pixel cluster can be determined in a number of ways. The cluster of pixels in **(E–H)** can be grouped as one, two, three, or four individuals, with **(I)** the optimal grouping being three individuals based on some quantifiable measure. When overlaps are large or body shapes are non-rigid, other methods using past and future dynamics are more suitable (see main text). **(J)** More complex contours can precisely map the pose of individuals, such as swimming in *Caenorhabditis elegans* [19] (Movie S2 in the supplementary material online), wing positioning in *Drosophila* [8] (Movie S14 in the supplementary material online), or body posturing of mice during social interactions [28] (Movie S11 in the supplementary material online). **(K)** Fingerprinting allows for maintenance of identities through time by analyzing the complete image structure, often using differences between individuals that are undetectable to the human eye, such as these zebra fish [20] (Movie S5 in the supplementary material online). Once individuals are detected and identified, their positions are linked across frames to form trajectories. **(L)** This could be a single individual in a 2D landscape [27], **(M)** a single individual in a 3D landscape (shown here with some habitat complexity) [18] (Movie S6 in the supplementary material online), **(N)** multiple individuals in a simple 2D landscape [27] (Movie S1 in the supplementary material online), or **(O)** multiple individuals in a 3D landscape (Movie S8 in the supplementary material online). **(P)** Trajectories throughout complex habitat can also be obtained, such as this woodlice navigating for 1 h between two habitat patches connected by a dispersal corridor (A.I. Dell, unpublished). See Acknowledgments for credits and permissions.

## Box 4. Analysis of tracking data

Automation results in vast quantities of high-quality behavioral data, which not only makes data management a key consideration but also presents major challenges in crystalizing this information into tractable and meaningful statistics. This problem is not unique, and it is possible to borrow data management and analysis techniques developed within other 'big data' fields, such as molecular biology and bioinformatics.

The most basic output from automated tracking is the coordinates of the center of body mass of one or more individuals through time (Box 3). Converting from pixel values into real world coordinates is often not as simple as using a pixel-to-distance scale factor, because even in situations with little depth variation the effects of perspective and foreshortening can be important. These issues can be readily overcome with standard photogrammetric techniques [91], or if the filming arena is not flat by integrating a 3D model of the surface into these calculations [34]. Once coordinates (and pose estimates if available) are produced, then even very simple analysis can address basic ecological questions such as where and how animals behave and interact [4,8] (Figure IA–C).

Higher order patterns in position and pose data can identify individual or between-individual behaviors [68] (Figure ID,E). Investigation of the relative position and behavior between individuals,

such as conspecifics [20,38,62,82] (Movie S1, Movie S2, Movie S5, Movie S7, Movie S8, and Movie S11 in the supplementary material online) or interacting predators and prey [4] (Movie S3 and Movie S9 in the supplementary material online), provide significant insights into mechanisms underlying the strength and outcome of ecological interactions, and the role of the physical environment [1] (Movie S10 in the supplementary material online). For instance, how much time do animals spend grooming, courting, searching, or chasing, and where and when do they perform these actions, and for interacting individuals, what are the relationships between their subjective sizes, body angles, and relative directions of motion [92]? Ultimately, automated behavioral analysis is limited only by our ability to quantitatively define behavior, or in our ability to develop machine-learning algorithms that can do this for us [73]. One popular procedure is to have a human identify behaviors in video, such as grooming or eating, without defining them. Given several examples of each behavior, a computer algorithm can learn distinguishing data features, creating a classifier or internal model of each behavior. This classifier can then be applied to new datasets, including new individuals and potentially other species. User-friendly, automated tools, such as JAABA [73], make these highly technical analyses accessible to non-experts.
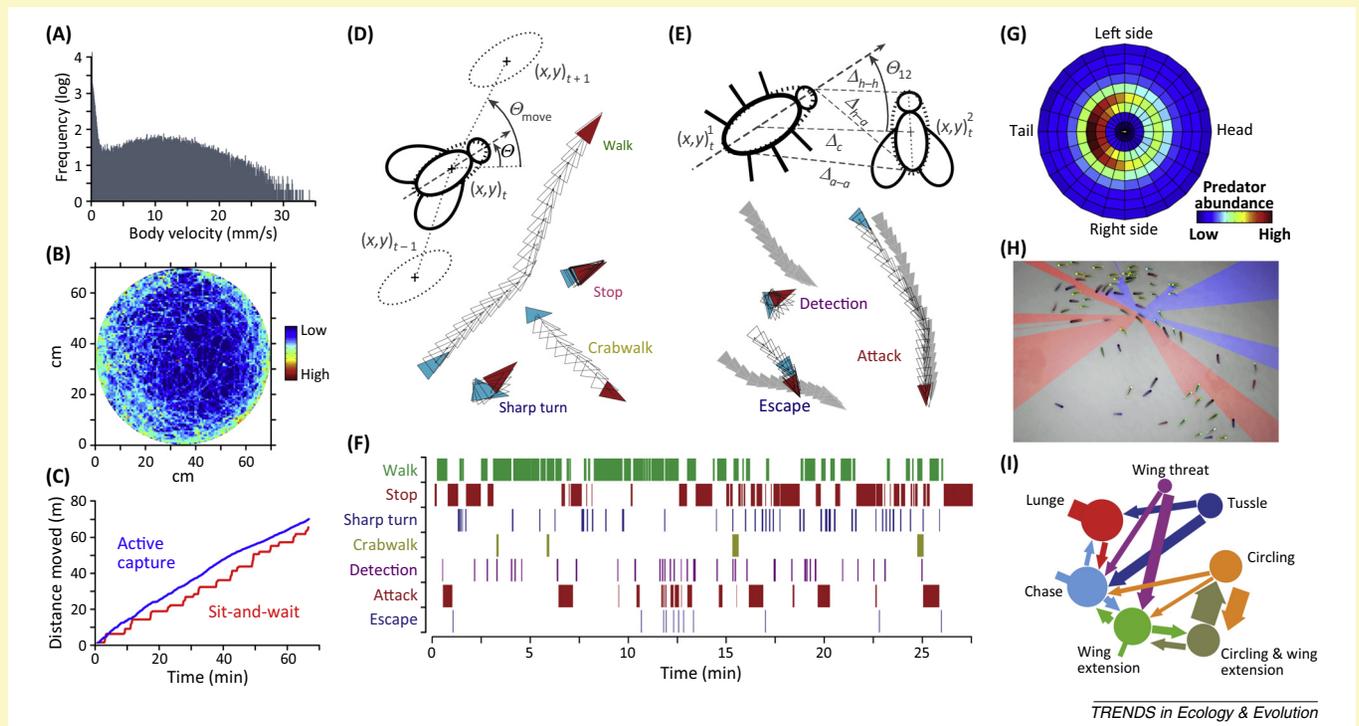


Figure I. The final step in automated image-based tracking is analysis, where position and pose data are analyzed to understand relevant biological, and ecological, patterns and processes. Simple statistics of positional data for individuals include (A) frequency distributions of body velocity, (B) location intensity maps revealing where throughout the landscape individuals spend their time, and (C) distance time plots, which can highlight foraging strategies employed by predators. (A, B) are data from an adult wolf spider (*Lycosa*) moving around a 65 cm diameter circular arena for 1 h, and (C) is a wolf spider (*Lycosa*, red line) and a centipede (*Lithobius*, blue line) under the same conditions (A.I. Dell, unpublished). This behavioral data can be automatically condensed into simpler, interpretable categories that characterize real behavioral phenotypes, either for (D) single individuals or (E) between individuals. The top panel in (D, E) show a contour model of the individual/s, together with the quantities (or traits) that can be measured, such as (D) a fly or (E) an insect predator feeding on a fly [8,27]. Symbol definitions are: $x,y$, spatial coordinates of the midpoint of the individual; $t$, time; $\Theta$, orientation; $\Theta_{move}$, moving direction; $\Delta$, distance; $a$, abdomen; $h$, head; $c$, midpoint of the animal [8]. The bottom panel (D,E) show some example behaviors that can be derived from these basic quantities [27]. Blue and red triangles are plotted at the start and end of each behavior example. For (E), the position of the non-target individual is plotted in gray. (F) From these behavioral assignments, time budgets can reveal how individuals spend their time or the components of ecological interactions (A.I. Dell, unpublished data). Analysis of between-individual position data can highlight patterns in the spatial arrangement of individuals within groups and aid in understanding collective behavior. Panel (G) shows the angular density of predators around a representative focal prey, with high predation density behind the prey revealing that piscivorous predators tend to attack from behind their prey [4], and panel (H) shows a reconstruction of the visual field of an individual fish embedded in the center of a large school, which provides novel insight into information flows within animal groups [26]. Behavioral information can also reveal the temporal linkage between complex behaviors, such as (I), the transitional behavioral graph for social interactions in *Drosophila* [8]. See Acknowledgments for credits and permissions.
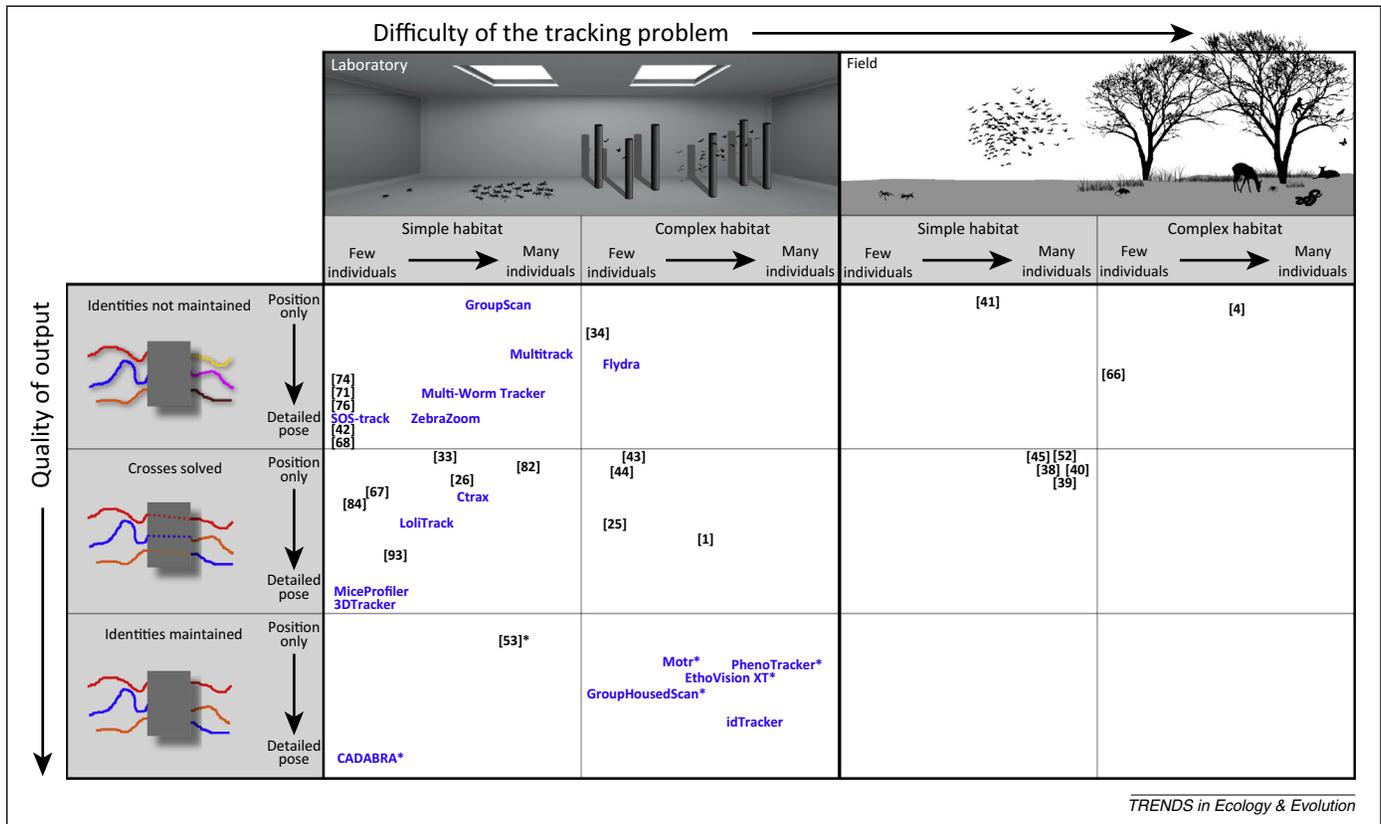
**Figure 2**. A number of key automated image-based tracking systems and studies relevant to ecologists, ranging on the horizontal axis from low numbers of individuals interacting in a simple landscape in the laboratory (left panel) to a natural system within a complex biological and physical landscape (right panel). The vertical axis separates studies and tracking systems by the type of output provided, specifically whether and how identity is maintained and whether position or detailed pose are tracked. In some cases, positions of labels have been slightly moved for visual clarity. Tracking systems can cover multiple categories, but are only shown once as tracking in simpler habitats with smaller numbers of individuals and less detailed pose will almost always be possible. See Table S1 in the supplementary material online for more details about each tracking system. Publications that use any of the tracking systems in Table S1 are not listed separately here. References are denoted in square brackets [1,4,25,26,33,34,38–45,52,53,66–68,71,74,76,82,84,93]. * Denotes the use of bio-logging or marking.

although this is computationally challenging and certainly some time away from use in tracking multiple moving targets. Other more recent and promising developments in hardware are single point 3D imaging technologies.

RGB-D (red, green, blue, depth) cameras, such as the Microsoft Kinect (www.microsoft.com/en-us/kinectforwindows/), achieve this by combining a color video camera either with an infrared projector to create a split infrared

---

**Box 5. Call to developers: the ideal automated image-based tracking system for ecologists**

Ecologists have a clear need for systems that can robustly and quickly gather large amounts of precise data on the behavior of multiple individuals from a diversity of species. Development of an automated image-based tracking system like this will require integrating disparate pieces of technology and software into a coherent, and user-friendly, package. This appears possible in the near future because of the continued increase in power, speed, and capabilities of hardware and software technologies. Any system that does this should have a number of key features.

(i) Simple to use. Automated image-based tracking systems involve a complex integration of software and hardware (see Figure 1 in main text), but many ecologists have insufficient funding to purchase expensive imaging and computational equipment, nor the technical expertise required to use them. Although there is some responsibility for ecologists to acquire these skills, tracking system development is ultimately the role of researchers and engineers with expertise in computer vision and informatics.

(ii) Marking should be unnecessary because this is time consuming, requires capture of individuals, and can alter their behavior and how other organisms interact with them.

(iii) Flexible enough that it can successfully track in a diverse range of experimental conditions, and individuals that vary greatly in their size, shape, and patterns of behavior.

(iv) Tracking large numbers of individuals is necessary because ecological systems are often characterized by high densities.

(v) Ecological systems are naturally embedded within a diverse range of environmental contexts, so it is essential that automated tracking systems function within diverse and complex habitats. The development of tracking software able to isolate animals from complex backgrounds, together with increasing automation of behavioral analysis, means the capabilities of automated image-based tracking for field use will only increase in the coming years.

(vi) The system must overcome the significant data storage and data management issues that inevitably arise when tracking at larger spatial and temporal scales. This can be partly overcome by real-time tracking, which reduces the need for processing and storing large amounts of digital data.

(vii) A single image point (i.e., camera) would be preferred over requiring multiple cameras, which can be difficult to integrate into a coherent system and can introduce disturbance effects that alter natural behavior.

(viii) Be mostly automated so that tracking and analysis is quick and consistent, including identifying individuals and their interactions, and quantifying their behavior in meaningful ways. However, it should still provide flexibility in the ways in which users can extract data, including at all levels of data acquisition, tracking, and analysis.

laser light field from which depth can be obtained (first generation) or by using time of arrival of the photons themselves (second generation). Light-field video cameras provide another promising technology (Box 2), where composite optics are used to simultaneously capture images focused at multiple distances from the lens, thus allowing for *post-hoc* selection of focus and ultimately 3D construction of the scene. As in 2D, multiple 3D imaging cameras can be employed simultaneously to provide additional resolution and to cope with occlusions [29].

*Into the field…*
Ecological systems are naturally embedded within environmental landscapes that are considerably more complex than laboratory arenas, such as streams, coral reefs, or the forest floor (Movie S13 in the supplementary material online). Although salient questions can be addressed in the laboratory, it is critical that tracking can be undertaken in the field because environmental drivers in their natural context – such as light, temperature, physical habitat, and spatial dimensionality – have profound influences on behavior and thus ecosystem organization [1,3,48,49] (Box 1). Many of the techniques that enable automated image-based tracking in the field are similar to those that enable 3D tracking, such as multiple cameras and single point 3D imaging devices (see above).

One of the primary constraints in the field is the ability to distinguish individuals within each image from the background (Box 2), which often varies unpredictably owing to such factors as wind, water, and sunlight. The simplest method to track in complex environmental landscapes is to employ an imaging method that provides clear contrast between the organisms and the background. The growing number of imaging technologies now available (Box 2) means there is a corresponding growing range of environmental contexts within which individuals can be tracked. For example, it is now straightforward to image independently of visible light (Box 2), meaning that tracking is no longer limited to environments with sufficient and homogeneous visible light [1,4,50–52] (Box 2). Another alternative is to use computer vision technologies that detect animals even when their color pattern is statistically indistinguishable from the background, based, for example, on their shape or movement [21]. Finally, it is possible to mark individuals [53] or integrate with other tracking methods such as bio-logging – combining the robustness of bio-loggers for detecting individuals in complex habitat with the high spatiotemporal resolution of imaging [54].

Physical structure, such as plant cover or soil, is more difficult to track within because it increases the number of occlusions. Fingerprinting allows the addition of habitat structure without increasing assignment errors, as identities are recovered following occlusion [20,21] (Movie S5 in the supplementary material online). Again, use of multiple cameras (Movie S6 in the supplementary material online), marking [53], or integration with bio-logging [54] can enable tracking and identity maintenance during or after occlusions. Some imaging methods can even pass through physical structure, revealing the position of organisms either behind the structure or embedded within it, thus removing the problem of occlusions from physical habitat altogether. The behavior of small invertebrates within soil, for example, has been successfully quantified using high-resolution X-ray microtomography, which works because biological tissue attenuates X-rays less than the surrounding soil matrix [55] (Box 2). Acoustic imaging (sonar) also permits imaging through relatively complex habitats, and is especially effective in aquatic environments. Modern high-resolution sonar has allowed biologists to investigate predator–prey interactions in habitats that would be impossible with other imaging methods [4] (Movie S9 in the supplementary material online) (Box 2). Additional technologies for imaging through complex physical habitats are on the horizon [56], although these are still probably years away from being successfully integrated into automated tracking systems.

For many ecological questions it is necessary to obtain quantitative information about the environmental landscape, which can be integrated with tracking data to understand how the environment influences behavior [1,5,17,57] (Box 1). Remote quantification of the environment is a key advantage of imaging over bio-logging, which only provides environmental information in the immediate vicinity of the individual to which the logger is attached. Remote quantification of the environment can easily be accomplished by imaging in the appropriate sensory regime, such as optical video cameras for quantifying light conditions and thermal cameras for quantifying the thermal landscapes. Methods for quantifying the physical structure of 3D landscapes are rapidly advancing [58–60] and can be used for rendering features of natural habitats, such as trees or streams. When combined with behavioral data, this environmental information should allow biologists to represent an animal's cognitive map of its environment, and thus understand the relationship between behavior and fitness [61]. This should be especially rewarding when combined with methods that reconstruct the sensory fields of individuals, providing knowledge about the sensory information on which animals base decisions [26].

## Quality of the tracking output
Ideally, the final output of tracking is the trajectory of each individual, spanning the entire image sequence and including detailed information about body posture and positioning of appendages. Realistically, however, this is a difficult outcome to obtain. We recognize two primary factors determining the quality of the tracking output: (i) how well identities are maintained throughout the image sequence, and (ii) whether only the midpoint or the detailed body posture of each individual is tracked.

*Identity maintenance*
We recognize three broad categories for how well identity is maintained by automated tracking systems (vertical axis in Figure 2). In the first category, identity is not maintained following occlusion, and instead new trajectories are produced each time a new individual is recognized (top row in Figure 2). In the second category, algorithms link identities across occlusions, based, for example, on the predicted movement of individuals (middle row in Figure 2).

When the number of individuals is not constant, as in the field, or if occlusions are too complex to link trajectories, the output can become similar to the first category. To prevent identity switch errors, the researcher must manually review uncertain events in the tracking data and make appropriate corrections (see above) [27,28,33,62] (Movie S1, Movie S10, and Movie S11 in the supplementary material online).

In the third scenario, each trajectory always belongs to a single individual (bottom row in Figure 2), similar to biologging. Here, individual organisms are recognized in each image, so that following occlusion correct identities are always maintained and identity errors never propagate. This often involves application of artificial markings; however, natural variation in the morphology of individuals can also be used to maintain identities throughout image sequences, even following occlusion (Table S1 in the supplementary material online). The simplest method involves using very general traits to identify individuals, such as body size or body shape. General traits can be sufficient for maintaining identities at low densities or when individuals vary greatly in size or shape, but in many other instances in ecology individuals are likely to be similarly sized or shaped. 'Fingerprinting' uses a more comprehensive set of traits to recognize individuals, even when individuals are indistinguishable to the human eye [20,21] (Table S1 and Movie S5 in the supplementary material online). Fingerprinting is currently limited to small numbers of individuals and controlled laboratory conditions, and although these limitations will be reduced as image quality increases and fingerprinting is combined with other segmentation and recognition methods [21,63–68], fingerprinting will always be limited for very large groups owing to inevitable 'overlaps' in fingerprints.

*Spatiotemporal position or detailed pose?*
Knowledge about the spatiotemporal position of organisms relative to the environmental landscape, or to each other, is sufficient for many questions in ecology [1,3,4,33,53] (Box 1). For other questions, such as those about the mechanics of locomotion or ecological interactions, it is necessary to know about the positioning of appendages or other specific points along an individual's body [8,17,19,28,29,69] (Movie S2, Movie S11, Movie S14, and Movie S22 in the supplementary material online) (Box 1). Estimating the center of body mass (position) is much simpler than detecting the detailed body posture and position of appendages (pose). Obtaining detailed pose information is not only a technically difficult computer vision problem but also requires higher spatial resolutions which brings with it associated costs, such as reduced spatial extent of imaging and increased data management and processing requirements [17].

Many tracking systems automatically identify individuals in images by fitting contour models to foreground pixels, once they have been isolated from the background (Box 3). These contour models can be simple, providing position and sometimes orientation [27] (Movie S1 and Movie S3 in the supplementary material online), or they can be more complex and thus provide detailed information about body posture or the position of legs, wings, or tails [8,19,28,29,68,70] (Movie S2, Movie S11, Movie S14, Movie S18, and Movie S22 in the supplementary material online) (Box 3). Automated tracking systems that use approaches other than contour fitting [17,20,68] (Movie S5 and Movie S17 in the supplementary material online) are more generalizable to different sized and shaped organisms, a key feature ecologists will ultimately require from tracking systems.

## Automated behavioral analysis

In addition to estimating the position and pose of individuals, automated tracking systems can also analyze individual and between-individual behaviors (behavior between individuals is analyzed in much the same way as for single individuals, except the context becomes behavioral correlations between individuals) (Box 4). Automatically annotating behavior produces large quantities of consistently defined and highly resolved behavioral data, providing biologists with unprecedented power to quantitatively understand general mechanisms and principles underlying behavior [68,71] (Box 1). Automated behavioral analysis is possible with trajectory information alone, such as differentiating between an individual being stationary, walking, or running, or with more detailed pose information, such as head position, contour shape, or appendage position [8,62,68,72] (Movie S2, Movie S14, and Movie S11 in the supplementary material online).

Identification of behavior into categories is a long-standing tradition in ethology and can be achieved by automated behavioral phenotyping, where the complex trajectory and pose data output by automated tracking is categorized into simpler, interpretable categories that best characterize biologically and ecologically relevant behavioral phenotypes (Box 4). Automated behavioral phenotyping can be undertaken with either supervised or unsupervised machine learning, with the core goal of both approaches being to condense the very rich and complex trajectory and pose data into a simpler form that is biologically and ecologically relevant (Box 4).

In supervised learning, the human expert identifies and categorizes patterns in the data by informing the software of categorical behavior annotations [8,15,71,73] (Movie S11, Movie S12, and Movie S14 in the supplementary material online). For example, 'wing grooming' could represent when a fly rubs one or both metathoracic legs over the top or the underside of the wing(s) [73]. Categorical annotations like these simultaneously take into account many different features of the trajectory and pose data, and result in categories that are generally easily interpretable by biologists and often have a clear physiological or ecological significance. Besides allowing higher throughput that manual annotation, the subjective *a priori* definitions of behavior chosen by a researcher can be expressed in precise mathematical terms and, once behaviors are defined, the analysis of any dataset is fully repeatable.

Unsupervised learning methods, by contrast, apply computational techniques to the raw data themselves to reveal what degrees of freedom are relevant in the data and then automatically detect any stereotyped patterns (Movie S12 in the supplementary material online) [68]. These behavioral patterns might, or might not, be evident to

human observers. Unsupervised techniques offer the advantage of decreased subjectivity, and increased throughput, repeatability, and the chance of finding rare behaviors [68,74,75]. Unsupervised methods use statistical methods to reduce the dimensionality of trajectory and pose data, but of course these statistical methods themselves depend on mathematical, human-generated assumptions inherent in the algorithms. It is therefore critical to compare the output of these unsupervised methods to what biologists already know about the inherent structure of behavior. As unsupervised methods become increasingly powerful and more objective [68], they will become an important development for community ecologists, who often study many different species which otherwise would require manual behavioral labeling across taxa, which is time consuming and prone to inaccuracies.

## Concluding remarks and future directions

Automated image-based tracking provides detailed information about the behavior of individuals at local scales, such as how they move, with which other individuals they interact, the sensory information available to them, and the role of internal and environmental drivers in shaping their behavior. This information should prove integral in mapping linkages among genes, brain function, behavior, and species interactions (thus linking molecular biology, neurobiology, and ecology). Although the development of automated image-based tracking has been primarily laboratory-based, focused on model organisms in controlled conditions, studies using this method now exist from a diversity of taxa and habitats, including in the field (Figure 2, Box 1, Table S1 in the supplementary material online). Automation of data collection permits high levels of replication, substantially increasing the amount and quality of behavioral data available to biologists. For example, automated tracking of functional response experiments, which ecologists use to quantify species interaction strength, would allow researchers to run large numbers of replicate trials where the behavior of every predator and prey is known (A.I. Dell, unpublished), providing unparalleled power to uncover novel patterns and biological mechanisms.

Similar to all technologies, automated image-based tracking has its limitations. Only individuals within the imaged area can be tracked – meaning larger animals cannot be tracked over their entire home range – and tracking in complex physical habitats or at high densities is difficult. In addition, the storage and management issues that arise from the huge amounts of digital data that are easily produced by imaging must be addressed. Emerging methods, such as fingerprinting or real-time tracking, alleviate this problem somewhat [20], but more needs to be done. As limitations are overcome, and imaging and computational technologies advance, automated image-based tracking should become firmly established as a powerful tool for quantitative research in ecology [9]. These methods are already providing conceptual advances on diverse topics such as predator–prey interactions, collective behavior of animal groups, social hierarchy, and population dynamics, and their continued application will only broaden this list of topics (Box 1; Table S1 in the supplementary material online). Eventually, automated tracking should influence the field of ecology similarly to how it is influencing the genetic and behavioral sciences, allowing ecologists to uncover mechanisms and principles that shape ecological systems, leading to a more general and predictive science of community ecology with significant basic and applied benefits.

## Author contributions

A.I. Dell conceived the project and wrote the first draft of the paper; all other authors contributed significantly. A.I. Dell made final decisions concerning the inclusion or exclusion of any technique or vendor covered in this review.

## Disclaimer statement

L.P.J.J. Noldus is Director of Noldus Information Technology BV, developer of one of the commercial tracking systems described in this paper (Table S1 in the supplementary material online).

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.tree.2014.05.004.

## References

1 Berdahl, A. *et al.* (2013) Emergent sensing of complex environments by mobile animal groups. *Science* 339, 574–576
2 Simpson, S.J. *et al.* (2010) Modelling nutritional interactions: from individuals to communities. *Trends Ecol. Evol.* 25, 53–60
3 Pawar, S. *et al.* (2012) Dimensionality of consumer search space drives trophic interaction strengths. *Nature* 486, 485–489
4 Handegard, N.O. *et al.* (2012) The dynamics of coordinated group hunting and collective information transfer among schooling prey. *Curr. Biol.* 22, 1213–1217
5 Nathan, R. *et al.* (2008) A movement ecology paradigm for unifying organismal movement research. *Proc. Natl. Acad. Sci. U.S.A.* 105, 19052–19059
6 Altmann, J. (1974) Observational study of behavior: sampling methods. *Behaviour* 49, 227–267
7 Tinbergen, N. (1963) On aims and methods of ethology. *Zeitschr. Tierpsychol.* 20, 410–433
8 Dankert, H. *et al.* (2009) Automated monitoring and analysis of social behavior in *Drosophila*. *Nat. Methods* 6, 297–303
9 Rutz, C. and Bluff, L.A. (2008) Animal-borne imaging takes wing, or the dawn of 'wildlife video-tracking'. *Trends Ecol. Evol.* 23, 292–294 author reply 294

10 Cooke, S.J. *et al.* (2004) Biotelemetry: a mechanistic approach to ecology. *Trends Ecol. Evol.* 19, 334–343

11 Krause, J. *et al.* (2013) Reality mining of animal social systems. *Trends Ecol. Evol.* 28, 541–551

12 Elliott, J.P. *et al.* (1977) Prey capture by the African lion. *Can. J. Zool.* 55, 1811–1828

13 Bayne, B.L. and Scullard, C. (1978) Rates of feeding by Thais-(Nucella)-Lapillus (L). *J. Exp. Marine Biol. Ecol.* 32, 113–129

14 Anonymous (2007) Geneticist seeks engineer: must like flies and worms. *Nat. Methods* 4, 463

15 Balch, T. *et al.* (2006) How multirobot systems research will accelerate our understanding of social animal behavior. *Proc. IEEE* 94, 1445–1463

16 Noldus, L.P.J.J. *et al.* (2002) Computerised video tracking, movement analysis and behaviour recognition in insects. *Comput. Electron. Agric.* 35, 201–227

17 Gomez-Marin, A. *et al.* (2012) Automated tracking of animal posture and movement during exploration and sensory orientation behaviors. *PLoS ONE* 7, e41642

18 Straw, A.D. *et al.* (2011) Multi-camera real-time three-dimensional tracking of multiple flying animals. *J. R. Soc. Interface* 8, 395–409

19 Swierczek, N.A. *et al.* (2011) High-throughput behavioral analysis in *C. elegans*. *Nat. Methods* 8, 592–598

20 Pérez-Escudero, A. *et al.* (2014) idTracker: tracking individuals in a group by automatic identification of unmarked animals. *Nat. Methods* http://dx.doi.org/10.1038/nmeth.2994

21 Kuhl, H.S. and Burghardt, T. (2013) Animal biometrics: quantifying and detecting phenotypic appearance. *Trends Ecol. Evol.* 28, 432–441

22 Krause, J. *et al.* (2011) Interactive robots in experimental biology. *Trends Ecol. Evol.* 26, 369–375

23 Ofstad, T.A. *et al.* (2011) Visual place learning in *Drosophila melanogaster*. *Nature* 474, 204–207

24 Straw, A.D. *et al.* (2010) Visual control of altitude in flying *Drosophila*. *Curr. Biol.* 20, 1550–1556

25 Branson, K. and Belongie, S. (2005) Tracking multiple mouse contours (without too many samples). *IEEE Comput. Vis. Pattern Recogn.* 1, 1039–1046

26 Strandburg-Peshkin, A. *et al.* (2013) Visual sensory networks and effective information transfer in animal groups. *Curr. Biol.* 23, R709–R711

27 Branson, K. *et al.* (2009) High-throughput ethomics in large groups of *Drosophila*. *Nat. Methods* 6, 451–457

28 de Chaumont, F. *et al.* (2012) Computerized video analysis of social interactions in mice. *Nat. Methods* 9, 410–417

29 Matsumoto, J. *et al.* (2013) A 3D-video-based computerized analysis of social and sexual interactions in rats. *PLoS ONE* 8, e78460

30 Brose, U. *et al.* (2006) Consumer-resource body-size relationships in natural food webs. *Ecology* 87, 2411–2417

31 Ohayon, S. *et al.* (2013) Automated multi-day tracking of marked mice for the analysis of social behaviour. *J. Neurosci. Methods* 219, 10–19

32 Simon, J.C. and Dickinson, M.H. (2010) A new chamber for studying the behavior of *Drosophila*. *PLoS ONE* 5, e8793

33 Katz, Y. *et al.* (2011) Inferring the structure and dynamics of interactions in schooling fish. *Proc. Natl. Acad. Sci. U.S.A.* 108, 18720–18725

34 Robie, A.A. *et al.* (2010) Object preference by walking fruit flies, *Drosophila melanogaster*, is mediated by vision and graviperception. *J. Exp. Biol.* 213, 2494–2506

35 Lacey, E.S. and Carde, R.T. (2011) Activation, orientation and landing of female *Culex quinquefasciatus* in response to carbon dioxide and odour from human feet: 3-D flight analysis in a wind tunnel. *Med. Vet. Entomol.* 25, 94–103

36 Spitzen, J. *et al.* (2013) A 3D analysis of flight behavior of *Anopheles gambiae sensu stricto* malaria mosquitoes in response to human odor and heat. *PLoS ONE* 8, e62995

37 Cachat, J. *et al.* (2011) Three-dimensional neurophenotyping of adult zebrafish behavior. *PLoS ONE* 6, e17597

38 Ballerini, M. *et al.* (2008) Interaction ruling animal collective behavior depends on topological rather than metric distance: evidence from a field study. *Proc. Natl. Acad. Sci. U.S.A.* 105, 1232–1237

39 Attanasi, A. *et al.* (2013) *Tracking in three dimensions via multi-path branching*, Cornell University Library, arXiv:1305.1495

40 Attanasi, A. *et al.* (2013) *Superfluid transport of information in turning flocks of starlings*, Cornell University Library, arXiv:1303.7097

41 Butail, S. *et al.* (2012) Reconstructing the flight kinematics of swarming and mating in wild mosquitoes. *J. R. Soc. Interface* 9, 2624–2638

42 Fontaine, E.I. *et al.* (2009) Wing and body motion during flight initiation in *Drosophila* revealed by automated visual tracking. *J. Exp. Biol.* 212, 1307–1323

43 Viscido, S.V. *et al.* (2004) Individual behavior and emergent properties of fish schools: a comparison of observation and theory. *Marine Ecol. Prog. Ser.* 273, 239–249

44 Veeraraghavan, A. *et al.* (2006) Motion based correspondence for 3D tracking of multiple dim objects. In *ICASSP 2006 Proceedings: 2006 IEEE International Conference on Acoustics, Speech and Signal Processing (Vol. 2)*. IEEE

45 Attanasi, A. *et al.* (2013) *Wild swarms of midges linger at the edge of an ordering phase transition*, Cornell University Library, arXiv:1307.5631

46 Kanbara, M. *et al.* (2006) 3D scene reconstruction from reflection images in a spherical mirror. In *In 18th International Conference on Pattern Recognition, ICPR 2006 (Vol.) 4*. pp. 874–879, IEEE

47 Chen, Z.H. *et al.* (2011) Single-view reconstruction from an unknown spherical mirror. In *18th IEEE International Conference on Image Processing ICIP 2011*. pp. 2677–2680, IEEE

48 Kalinkat, G. *et al.* (2013) Habitat structure alters top-down control in litter communities. *Oecologia* 172, 877–887

49 Dell, A.I. *et al.* (2011) Systematic variation in the temperature dependence of physiological and ecological traits. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10591–10596

50 Hristov, N.I. *et al.* (2008) Applications of thermal infrared imaging for research in aeroecology. *Integr. Comp. Biol.* 48, 50–59

51 Chiu, C. *et al.* (2010) Effects of competitive prey capture on flight behavior and sonar beam pattern in paired big brown bats, *Eptesicus fuscus*. *J. Exp. Biol.* 213, 3348–3356

52 Betke, M. *et al.* (2008) Thermal imaging reveals significantly smaller Brazilian free-tailed bat colonies than previously estimated. *J. Mammal.* 89, 18–24

53 Mersch, D.P. *et al.* (2013) Tracking individuals shows spatial fidelity is a key regulator of ant social organization. *Science* 340, 1090–1093

54 Weissbrod, A. *et al.* (2013) Automated long-term tracking and social behavioural phenotyping of animal colonies within a semi-natural environment. *Nat. Commun.* 4, 2018

55 Johnson, S.N. *et al.* (2007) Non-invasive techniques for investigating and modelling root-feeding insects in managed and natural systems. *Agric. Forest Entomol.* 9, 39–46

56 Bertolotti, J. *et al.* (2012) Non-invasive imaging through opaque scattering layers. *Nature* 491, 232–234

57 Gomez-Marin, A. *et al.* (2011) Active sampling and decision making in *Drosophila* chemotaxis. *Nat. Commun.* 2, 441

58 Izadi, S. *et al.* (2011) KinectFusion: real-time 3D reconstruction and interaction using a moving depth camera. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*. pp. 559–568, ACM

59 Newcombe, R.A. *et al.* (2011) KinectFusion: real-time dense surface mapping and tracking. In *10th IEEE International Symposium on Mixed and Augmented Reality*. pp. 127–136, IEEE

60 Yang, X. *et al.* (2013) Three-dimensional forest reconstruction and structural parameter retrievals using a terrestrial full-waveform lidar instrument (Echidna®). *Remote Sens. Environ.* 135, 36–51

61 Cagnacci, F. *et al.* (2010) Animal ecology meets GPS-based radiotelemetry: a perfect storm of opportunities and challenges. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* 365, 2157–2162

62 Buhl, J. *et al.* (2006) From disorder to order in marching locusts. *Science* 312, 1402–1406

63 Fergus, R. *et al.* (2003) Object class recognition by unsupervised scale-invariant learning. In *Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Vol. II)*. pp. 264–271, IEEE

64 Dollár, P. *et al.* (2010) The fastest pedestrian detector in the west. In *British Machine Vision Conference*. (Labrosse, F. *et al.*, eds), pp. 68.1–68.11. BMVA Press

65 Burghardt, T. and Calic, J. (2006) Real-time face detection and tracking of animals. In *8th Seminar on Neural Network Applications in Electrical Engineering, NEUREL 2006*. pp. 27–32, IEEE

66 Hannuna, S.L. *et al.* (2005) Identifying quadruped gait in wildlife video. In *IEEE International Conference on Image Processing (Vol. 1), ICIP 2005*. pp. 713–716, IEEE

67 Burgos-Artizzu, X.P. *et al.* (2012) Social behavior recognition in continuous video. In *2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 1322–1329, IEEE

68 Berman, G.J. *et al.* (2013) *Mapping the structure of drosophilid behavior*, Cornell University Library, arXiv:1310.4249

69 Bender, J.A. *et al.* (2011) Kinematic and behavioral evidence for a distinction between trotting and ambling gaits in the cockroach *Blaberus discoidalis*. *J. Exp. Biol.* 214, 2057–2064

70 Mirat, O. *et al.* (2013) ZebraZoom: an automated program for high-throughput behavioral analysis and categorization. *Front. Neural Circuits* 7, 107

71 van Dam, E.A. *et al.* (2013) An automated system for the recognition of various specific rat behaviours. *J. Neurosci. Methods* 218, 214–224

72 Dollár, P. *et al.* (2010) Cascaded pose regression. In *2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 1078–1085, IEEE

73 Kabra, M. *et al.* (2013) JAABA: interactive machine learning for automatic annotation of animal behavior. *Nat. Methods* 10, 64–67

74 Brown, A.E. *et al.* (2013) A dictionary of behavioral motifs reveals clusters of genes affecting *Caenorhabditis elegans* locomotion. *Proc. Natl. Acad. Sci. U.S.A.* 110, 791–796

75 Braun, E. *et al.* (2010) Identifying prototypical components in behaviour using clustering algorithms. *PLoS ONE* 5, e9361

76 Fontaine, E. *et al.* (2008) Automated visual tracking for studying the ontogeny of zebrafish swimming. *J. Exp. Biol.* 211, 1305–1316

77 Censi, A. *et al.* (2013) Discriminating external and internal causes for heading changes in freely flying *Drosophila*. *PLoS Comput. Biol.* 9, e1002891

78 Maimon, G. *et al.* (2008) A simple vision-based algorithm for decision making in flying *Drosophila*. *Curr. Biol.* 18, 464–470

79 Pizzo, A.B. *et al.* (2013) The membrane raft protein Flotillin-1 is essential in dopamine neurons for amphetamine-induced behavior in *Drosophila*. *Mol. Psychiatry* 18, 824–833

80 van Breugel, F. and Dickinson, M.H. (2012) The visual control of landing and obstacle avoidance in the fruit fly *Drosophila melanogaster*. *J. Exp. Biol.* 215, 1783–1798

81 Wu, Z. *et al.* (2011) A combinatorial semaphorin code instructs the initial steps of sensory circuit assembly in the *Drosophila* CNS. *Neuron* 70, 281–298

82 Tunstrom, K. *et al.* (2013) Collective states, multistability and transitional behavior in schooling fish. *PLoS Comput. Biol.* 9, e1002915

83 Herbert-Read, J.E. *et al.* (2011) Inferring the rules of interaction of shoaling fish. *Proc. Natl. Acad. Sci. U.S.A.* 108, 18726–18731

84 Fan, J. *et al.* (2010) Automatic video-based analysis of animal behaviors. In *2010 17th IEEE International Conference on Image Processing (ICIP)*. pp. 1513–1516, IEEE

85 Wang, L. and Anderson, D.J. (2010) Identification of an aggression-promoting pheromone and its receptor neurons in *Drosophila*. *Nature* 463, 227–231

86 Brady, D.J. *et al.* (2012) Multiscale gigapixel photography. *Nature* 486, 386–389

87 Straw, A.D. and Dickinson, M.H. (2009) Motmot, an open-source toolkit for real-time video acquisition and analysis. *Source Code Biol. Med.* 4, 5

88 Stuurman, N. *et al.* (2007) Micro-Manager: open source software for light microscope imaging. *Microsc. Today* 15, 42–43

89 Bar-Shalom, Y. and Fortmann, T.E. (1988) *Tracking and Data Association*, Academic Press

90 Delcourt, J. *et al.* (2013) Video multitracking of fish behaviour: a synthesis and future perspectives. *Fish Fisheries* 14, 186–204

91 Hartley, R. and Zisserman, A. (2004) *Multiple View Geometry in Computer Vision*, Cambridge University Press

92 Zabala, F. *et al.* (2012) A simple strategy for detecting moving objects during locomotion revealed by animal–robot interactions. *Curr. Biol.* 22, 1344–1350

93 Butail, S. and Paley, D.A. (2012) Three-dimensional reconstruction of the fast-start swimming kinematics of densely schooling fish. *J. R. Soc. Interface* 9, 77–88