RESEARCH ARTICLE

Automatic detection and decoding of honey bee waggle dances

Fernando Wario*, Benjamin Wild, Raúl Rojas, Tim Landgraf

FB Mathematik und Informatik, Freie Universität Berlin, Berlin, Germany

* fernando.wario@fu-berlin.de

Abstract

The waggle dance is one of the most popular examples of animal communication. Forager bees direct their nestmates to profitable resources via a complex motor display. Essentially, the dance encodes the polar coordinates to the resource in the field. Unemployed foragers follow the dancer's movements and then search for the advertised spots in the field. Throughout the last decades, biologists have employed different techniques to measure key characteristics of the waggle dance and decode the information it conveys. Early techniques involved the use of protractors and stopwatches to measure the dance orientation and duration directly from the observation hive. Recent approaches employ digital video recordings and manual measurements on screen. However, manual approaches are very time-consuming. Most studies, therefore, regard only small numbers of animals in short periods of time. We have developed a system capable of automatically detecting, decoding and mapping communication dances in real-time. In this paper, we describe our recording setup, the image processing steps performed for dance detection and decoding and an algorithm to map dances to the field. The proposed system performs with a detection accuracy of 90.07%. The decoded waggle orientation has an average error of -2.92˚ (± 7.37˚), well within the range of human error. To evaluate and exemplify the system's performance, a group of bees was trained to an artificial feeder, and all dances in the colony were automatically detected, decoded and mapped. The system presented here is the first of this kind made publicly available, including source code and hardware specifications. We hope this will foster quantitative analyses of the honey bee waggle dance.

Introduction

The honey bee waggle dance is one of the most popular communication systems in the animal world. Forager bees move in a stereotypic pattern on the honeycomb to share the location of valuable resources with their nestmates [1–3]. Dances consist of waggle and return phases. During the waggle phase, the dancer vibrates her body from side to side while moving forward in a rather straight line on the vertical comb surface. Each waggle phase is followed by a return phase, during which the dancer circles back to the starting point of the waggle phase. Clockwise and counterclockwise return phases are alternated, such that the dancer describes a path

resembling the figure eight [1, 4–6]. The average orientation of successive waggle runs with respect to gravity approximates the angle between the advertised resource and the solar azimuth as seen from the hive Fig 1. The duration of the waggle run correlates with the distance between hive and resource [1, 7–9]. The resource's profitability is encoded in the dance tempo: valuable resources are signaled with shorter return runs, yielding a higher waggle production rate [10].

Unemployed foragers might become interested in a waggle dance, follow the dancer's movements and decode the information contained in the dance. Followers may then exit the hive to search for the communicated resource location in the field [2, 11–14]. Recruits that were able to locate the resource, once back in the hive, may also dance, thereby amplifying the collective foraging effort.

The study of the waggle dance as an abstract form of communication received great interest after it was first described by von Frisch [15]. Keeping bee colonies in special hives for observation is well-established. The complex dance behavior allows insights into many aspects of the honey bee biology and, even after seven decades, several research fields investigate the waggle dance communication system.

The dance essentially contains polar coordinates for a field location. Hence, waggle dances can be mapped back to the field [16]. The directional component relies on fixed
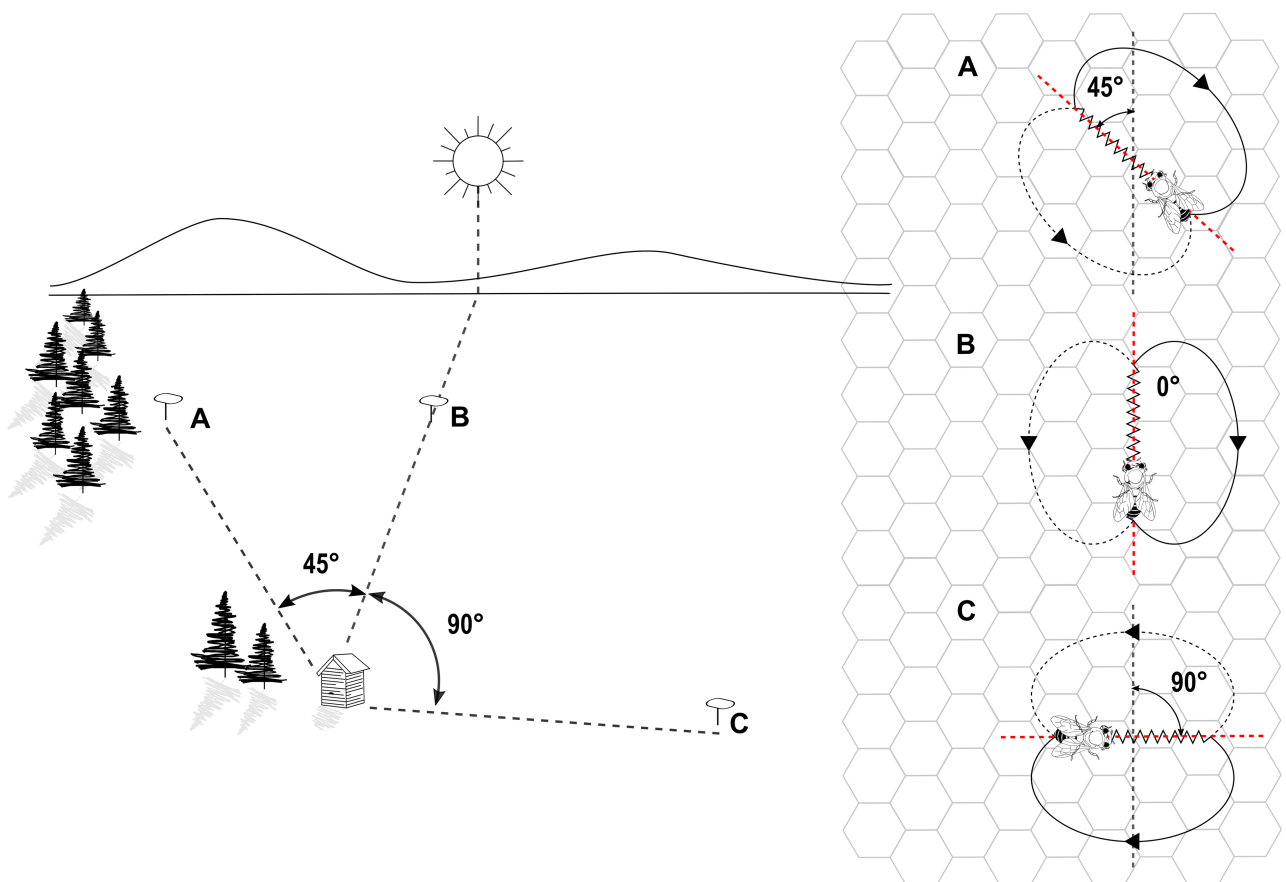

Fig 1. Correlation between waggle dance parameters and locations on the field. On the left, three food sources in the field located at A) 45° counterclockwise B) 0° and C) 90° clockwise, with respect to the azimuth. On the right, their representation through waggle dance paths on the surface of a vertical honeycomb.

reference systems such as gravity or the sun's azimuth and therefore is straightforward to compute from a dance observation. Honeybees integrate the optical flow they perceive along their foraging routes to gauge the distance they have flown [8]. Several factors affect the amount of optical flow perceived, such as wind (bees fly closer to the ground in with strong headwinds [17]) or the density of objects in the environment, such as vegetation or buildings. Honeybees calibrate their odometer to the environment before engaging in foraging activities. Hence, to convert waggle durations to feeder distances, our system requires calibration itself. To this end, bees must be trained to a number of sample locations with known distance. Assuming homogeneous object density in all directions, it may be sufficient to use a simple conversion factor obtained from the waggle durations observed in dances for a single feeding location.

This way, without tracking the foragers' flights, one can deduce the distribution of foragers in the environment by establishing the distribution of dance-communicated locations. Couvillon et al. [18] used this method to investigate how the decline in flower-rich areas affects honey bee foraging, while Balfour and Ratnieks [19] used it to find new opportunities for maximizing pollination of managed honey bee colonies. But mapping is not the only application of decoding bee dances. Theoretical biologists have studied the information content of the dance [20, 21] and the accuracy and precision with which bees represent spatial information through waggle dances [22, 23]. Landgraf and co-workers tracked honey bee dances in video recordings to build a motion model for a dancing honey bee robot [6, 24]. Studies on honey bee collective foraging also focus on the waggle dance [25], including studies that model their collective foraging [26, 27] and nest-site selection behavior [28–30]. In [31] we automatically decoded waggle dances as part of an integrated solution for the automatic long-term tracking of activity inside the hive.

Different techniques have been used over time to decode waggle dances. During the first decades that followed von Frisch's discovery, most of the dances were analyzed in real time, directly from the observation hive with the help of protractors and stopwatches [1, 2]. Throughout the last decade, the use of digital video has become ubiquitous to extract the encoded information on their computers. Digital video allows researchers to analyze dances frame by frame and extract their characteristics either manually using the screen as a virtual observation hive [23], or assisted by computer software [22]. Although digital video recordings allow measurements with higher accuracy and precision, decoding communication dances continued to be a manual and time-consuming task.

Multiple automatic and semi-automatic solutions have been proposed to simplify and accelerate the dance decoding process. A first group of solutions focused on mapping the bees' trajectories via tracking algorithms [32–34]. These trajectories might then be analyzed to extract specific features such as waggle run orientation and duration, using either a generic classifier trained on bee dances (see [35–37]) or methods based on hand-crafted features such as the specific spectral composition of the trajectory in a short window [6]. Although a method has been described by Feldman and Balch [38] that could potentially be an automatic detector and decoder of dances [38], its implementation has been limited to the automatic labeling of behaviors.

Here we propose a solution to detect and decode waggle dances automatically. Since all information known to be carried in the dances, can be inferred from the waggle run characteristics, our algorithm exclusively detects this portion of the dance directly from the video stream, avoiding a separate tracking stage. Our system detects 89.8% of all waggle runs with a false positive rate of only 5%. Compared to a human observer, the system extracts the

waggle orientation with an average error of -2.92˚ (± 7.37˚) well within the range of human error.

Materials and methods

Hive and recording setup

Our solution can be used either online with live streaming video or offline with recorded videos. The software requires a frame rate of approximately 100 Hz and a resolution of at least 1.5 pixels per millimeter. Thus QVGA resolution suffices to cover the whole surface of a "Deutsch-normal" frame (370 mm x 210 mm). The four frame corners are used as a reference to rectify distortions caused by skewed viewing angles or camera rotations; If the frame corners are not captured by the camera, it is necessary to consider other reference points and their relative coordinates. A basic setup configuration is depicted in Fig 2 and might serve as a template for the interested researcher.

In our experiments, we worked with a small colony of 2000 bees (*Apis mellifera carnica*). We used a one frame observation hive and one modified PS3eye camera per side. The camera is a low-cost model that offers frame rates up to 125 Hz at QVGA resolution (320 x 240 pixels) using an alternative driver (see S1 Text). For the lighting setup, it is necessary to use a constant light source, such as LEDs. Pulsed light sources, such as fluorescent lamps, may introduce flicker to the video, yielding suboptimal detection results. Our setup was illuminated by an array of infrared IR-LED clusters (840 nm wavelength). The entire structure was enveloped with a highly IR reflective foil with small embossments for light dispersion. The IR LED clusters pointed towards the foil to create a homogeneous ambient lighting and reduce reflections on the glass panes. The built in IR block of the PS3eye cameras had to be removed to make them IR sensitive.

Target features

The relation of site properties (distance and direction to the feeder) and dance properties (duration and angle) have been recorded via systematic experiments [1]. The following
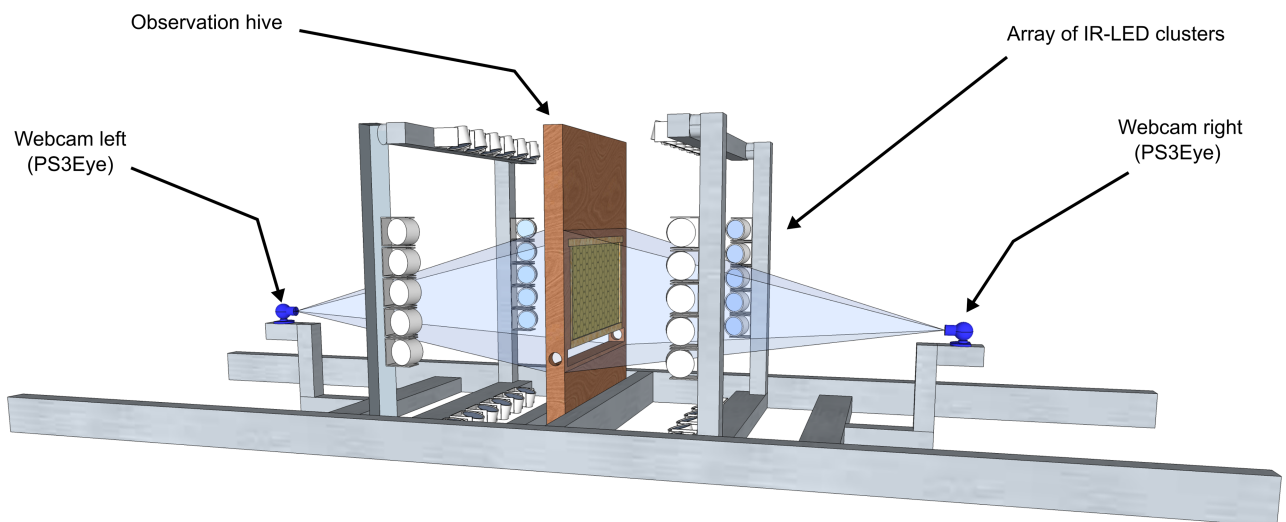

Fig 2. Recording setup. A basic recording setup consists of an observation hive, an array of LED clusters for illumination and one webcam per side.

equations will be used for an approximate inverse mapping of dance parameters to site properties.

$$r_R \sim f_d^{-1}(d_w).\tag{1}$$

$$\theta_R \sim atan2\left(\sum_{j=1}^{n} sin\alpha_{wj}, \sum_{j=1}^{n} cos\alpha_{wj}\right), n = 2k.\tag{2}$$

$$p_R \sim \frac{d_w}{d_r}.\tag{3}$$

Where r_R (Eq 1) is the distance between hive and resource. It is related to the average waggle run duration d_w through the function f_d that approximates the calibration curve [16]. Here, we use a simple linear mapping and use an empirically determined conversion factor. θ_R (Eq 2) is the angle between resource and solar azimuth (see Fig 1). It corresponds to the average orientation of the waggle runs with respect to the vertical, with an even number of consecutive runs to avoid errors due to the divergence angle [1, 4–6]. The resource's profitability p_R (Eq 3) is proportional to the ratio between average waggle run duration d_w and average return run duration d_r. Dances for high-quality resources contain shorter return runs than those for less profitable resources located at the same distance, hence yielding a higher p_R value [10].

From Eqs 1 to 3 it follows that to decode the information contained in a communication dance three measurements are required: average waggle run duration d_w, average orientation α_w, and average return run duration d_r. In contrast to some approaches that require tracing the dance path to then analyze it and extract its characteristics [32, 33], we propose an algorithm that directly analyzes video frames to obtain each waggle run's starting timestamp, duration, and angle. In our approach, return run durations are calculated as the time difference between the end of a waggle run and the beginning of the next one Fig 3.

Software modules

Our software consists of four modules that are executed in sequence, namely: attention module (*AM*), filter network (*FN*), waggle orientation module (*OM*) and mapping module (*MM*). The *AM* runs in real-time and stores small subregions of the video containing waggle-like activity. Later, false positives are filtered out using a convolutional neural network *FN*. The *OM* extracts the duration and angle of the waggle runs. Finally, the mapping module (*MM*) clusters waggle runs belonging to the same dances and maps them back to field coordinates. All modules can run offline on video recordings. Long observations, however, require large storage space. Therefore, we propose using the *AM* with a real-time camera input to reduce the amount of stored data drastically. A detailed description of all four modules is given in the following sections.
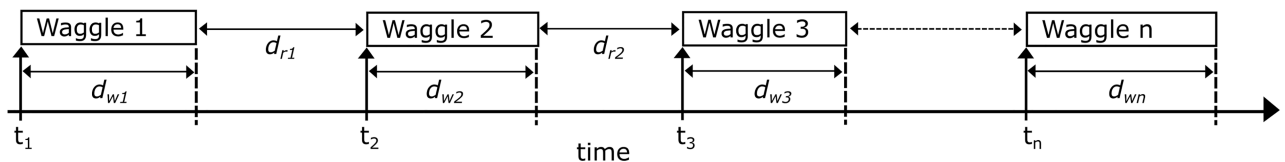

Fig 3. Fundamental parameters. Knowing the starting time (t_x) and duration (d_{wx}) for each waggle run, it is possible to calculate the return run durations as the time gaps between consecutive waggle runs.

Attention module (*AM*). The waggle frequency of dancing bees lies within a particular range we call the *waggle band*[34, 39]. Under consistent lighting conditions, the bee body is well discriminable from the background; therefore, the brightness dynamics of each pixel in the image originates from either honey bee activity or sensor noise. Thus, when a pixel intersects with a dancing bee, its intensity time-series is a function of the texture pattern on the bee and her motion dynamics. Indeed, by using a camera with low spatial resolution, bees appear as homogeneous ellipsoid blobs without surface texture. Thus the brightness of pixels that are crossed by waggling bees varies with the periodic waggle motion, and the frequency spectrum of that time series exhibits components in the *waggle band* or harmonics. To detect this general feature, we define a binary classifier here on referred to as *Dot Detector* (*DD*), each pixel position $[i, j]$ is associated to a *DD* D_{ij}. The *DD* analyzes the intensity evolution of the pixel within a sliding window of width b. For this purpose, the last b intensity values of each pixel are stored in a vector B, which at the time n can be described as $B_{ij}^n = [v_{ij}^{n-b+1}, v_{ij}^{n-b+2}, \ldots, v_{ij}^n]$, where v_{ij}^k is the intensity value of the pixel $[i, j]$ at time k. We calculate a score for each of these time series using a number of sinusoidal basis functions, in principle similar to the Discrete Fourier Transform [40]:

$$score(\bar{B}_{ij}^n, r) = \sum_{m=1}^{b} \left(\left(\bar{B}_{ij}^n(m) \cdot cos\left(2\pi r \frac{m}{s_r} \right) \right)^2 + \left(\bar{B}_{ij}^n(m) \cdot sin\left(2\pi r \frac{m}{s_r} \right) \right)^2 \right) \tag{4}$$

where $\bar{B}_{ij}^n$ is the normalized version of B_{ij}^n with $\bar{B}_{ij}^n \in \{-1, 1\}$ and $min_{ij}^n = min(B_{ij}^n)$ and $max_{ij}^n = max(B_{ij}^n)$, s_r is the video's sample rate (100 Hz), and $r \in [10, 16]$ are the frequencies in the waggle band. If at least one of the frequencies in the *waggle band* scores over a defined threshold t_h, D_{ij} is set to 1. After computing the scores, those D_{ij} set to 1 are clustered together following a hierarchical agglomerative clustering (*HAC*) approach [41], using as a metric the Euclidean distance between pixels and with a threshold d_{max1} set to half the body length of a honey bee. Clusters formed by less than c_{min1} *DDs* are discarded as noise-induced, and the centroids of the remaining clusters are regarded as positions of potential dancers.

Positions found during the clustering step are then used to detect waggle runs (*WR*). If positions detected in successive frames are located within a maximum distance d_{max2}, defined according to the average waggle forward velocity (see [6]), the positions are considered as belonging to the same *WR*. At each iteration new dancer positions are matched against open *WR* candidates, and either appended to a candidate or used as basis for a new one. A *WR* candidate can remain open up to g_{max2} frames without new detections being added. If no detections could be added it is closed. Only closed WR candidates with a minimum of c_{min2} detections are retrieved as *WRs*. Finally, coordinates of the potential dancer, along with 50 x 50 pixels image snippets of the *WR* sequence are stored to disk.

The operation of the *AM* can be seen as a three layers process summarized in the following points:

1. Layer 0, for each new frame I^n:

    (a) Update *DDs*' score vector.

    (b) Set to 1 *DDs* with spectrum components in the waggle band above t_h

2. Layer 1, detecting potential dancers:

    (a) Cluster together *DDs* potentially activated by the same dancer.

    (b) Filter out clusters with less than c_{min1} elements.

(c) Retrieve clusters' centroids as coordinates for potential dancers.

3. Layer 2, detecting waggle runs:

   (a) Create waggle run assumptions by concatenating dancers positions with a maximum Euclidean distance of d_{max2}.

   (b) Assumptions with a minimum of c_{min2} elements are considered as real *WR*.

Filtering with convolutional neural network (*FN*).   For long term observations we propose using the *AM* to filter relevant activity from a camera stream in real-time. This significantly reduces the disk space otherwise required to store full sized videos. Depending on the task at hand, it might e.g. be advisable to configure the module to never miss a dance. A higher sensitivity, however, might come with a higher number of false detections. For this use-case, we have trained a convolutional neural network that processes the sequence of 50 x 50 px images to discard non-waggles. The scalar output of the network is then thresholded to predict whether the input sequence contains a waggle dance. The network is a 3D convolutional network whose convolution and pooling layers are extended to the 3rd, i.e. temporal, dimension [42–45]. The network architecture, three convolutional and two fully connected layers, is rather simple but suffices for the filtering tasks (for details refer to S1 Fig).

The network was trained on 8239 manually labeled AM detections from two separate days. During training, subsequences consisting of 128 frames were randomly sampled from the detections for each mini-batch. Detections with less than 128 frames were padded with constant zeros. Twenty percent of the manually labeled data was reserved for validation. To reduce overfitting, the sequences were randomly flipped on the horizontal and vertical axes during training. We used the Adam optimizer [46] to train the network and achieved an accuracy of 90.07% on the validation set. This corresponds to a recall of 89.8% at 95% precision.

Orientation module (*OM*).   While the duration of a *WR* is estimated from the number of frames exported by the *AM*, its orientation is computed in a separate processing step here on referred as orientation module (*OM*), usually performed offline to keep computing resources free for detecting waggle runs. Dancing bees move particularly fast during waggle runs, throwing their body from side to side at a frequency of about 13 Hz [6]. Images resulting from subtracting consecutive video frames of waggling bees exhibit a characteristic pattern similar to a 2D Gabor filter, a positive peak next to a negative peak, whose orientation is aligned with the dancer's body Fig 4.

The Fourier transformation of the difference image provides a location-independent representation of the waggling event while preserving information regarding the dancer's orientation (Fig 4B). We make use of the Fourier slice theorem [47], which states that the Fourier transform of a projection of the original function onto a line at an angle α is just a slice through the Fourier transform at the same angle. Imagine a line orthogonal to the dancer's orientation. If we project the Gabor-like pattern onto this line, we obtain a clear sinusoidal pattern which appears as a strong pair of maxima in the Fourier space at the same angle. Not all difference images in a given image sequence exhibit the Gabor pattern, it only appears when the bee is quickly moving laterally. To get a robust estimate of the waggle orientation, we sum all Fourier transformed difference images Fig 5A and apply a bandpass filter Fig 5B to obtain the correct maxima locations Fig 5C. The bandpass filter is performed in the frequency domain by multiplying with a difference-of-Gaussians *DoG*. The radius of the ring needs to be tuned to the expected frequency of the sinusoidal in the 1D projection of the Gabor pattern. This frequency depends on the frame rate (in our case 100 Hz) and the image resolution (17 px/mm). With the lateral velocity of the bee (we used the descriptive statistics in [6]) one can compute the
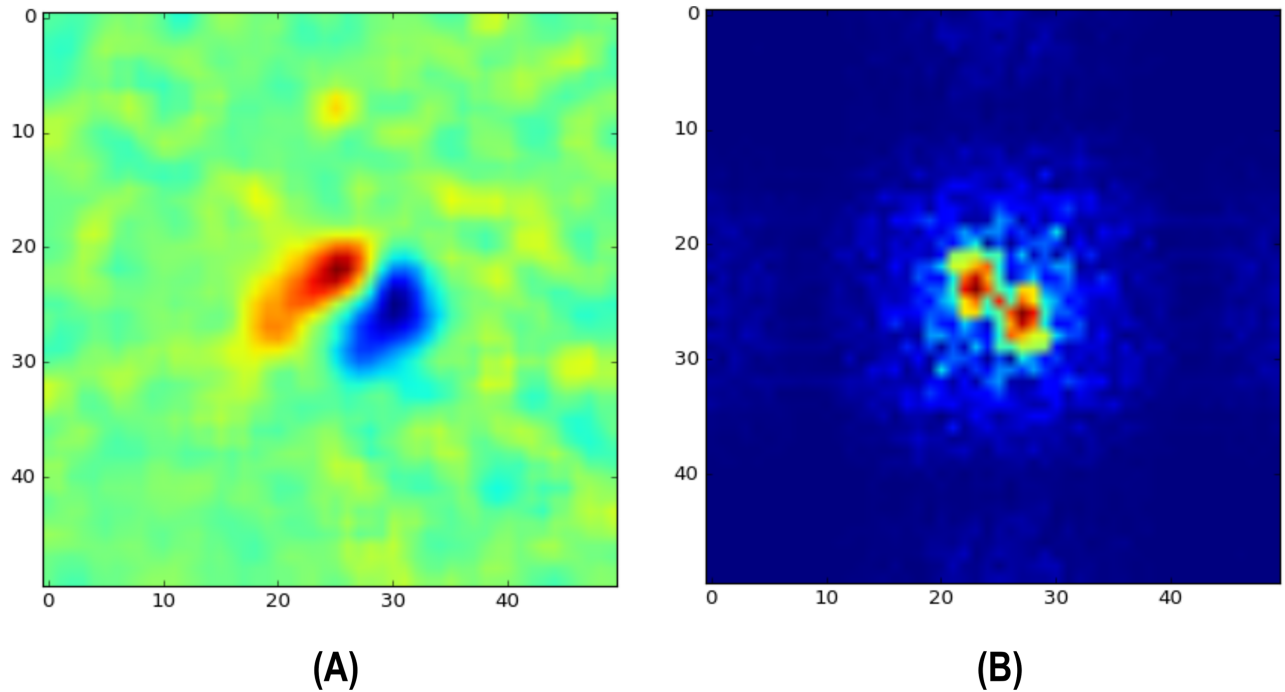
Fig 4. Difference image and its Fourier transformation. (A) The image resulting from subtracting consecutive video frames of waggling bees exhibits a characteristic Gabor filter-like pattern. (B) While the peak location varies in image space along with the dancer's position, its representation in the Fourier space is location-independent, showing distinctive peaks at frequencies related to the size and distance of the Gabor-like pattern.

displacement in pixels (5 to 7 px/frame). Using Eq 5 we can approximate the value of the expected frequency k:

$$k = \frac{I_{size}}{T} = \frac{I_{size}}{2x},$$ (5)
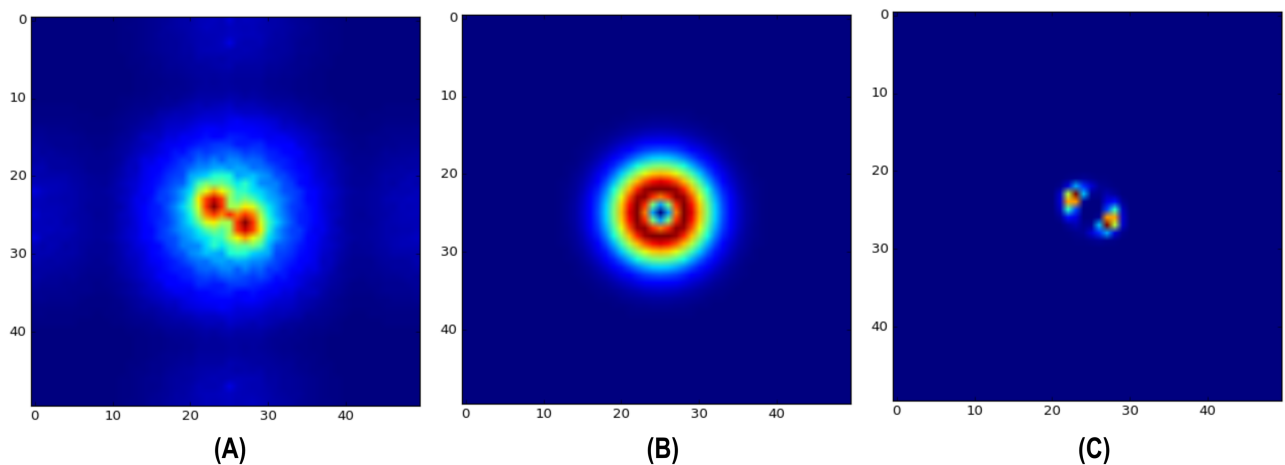
Fig 5. Filtering cumulative sum of difference images in the Fourier space. (A) The cumulative sum of the Fourier transformed difference images of a waggle run exhibit a strong pair of maxima in locations orthogonal to the dancer's orientation. (B) A *DoG* kernel of the Mexican hat type, properly adjusted to the waggle band, is used as a bandpass filter. (C) Bandpass filtering the cumulative sums emphasizes values within the frequencies of interest.

where I_{size} is the input image size (50 px in this case) and T is the period for the Gabor filter-like pattern or twice the bee's displacement between frames.

The orientation of the waggle run is obtained from the resulting image through Principal Component Analysis (*PCA*). However, the principal direction reflects the direction of the dancer's lateral movements, so it is necessary to add 90˚ to this direction to obtain the dancer's body axis. This axis represents two possible waggle directions. To disambiguate the alternatives, we process the dot detector positions extracted by the *AM*. Each of these image positions represents the average pixel position in which we found brightness changes in the waggle band. In a typical waggle sequence, these points trace roughly the path of the dancer. We average all *DD* positions of the first 10% of the waggle sequence and compute all *DD* positions relative to this average. We then search for the maximum values in the histogram of the orientation of all vectors and average their direction for a robust estimate of the main direction of the dot detector sequence. This direction is then used to disambiguate the two possible directions extracted by *PCA*.

Mapping module (*MM*).   Waggle dances encode polar coordinates for field locations. To map these coordinates back to the field we implemented a series of steps in what is here on referred as mapping module (*MM*). The *MM* reads the output of the *AM* and *OM*, essentially time, location, duration and orientation of each detected waggle run. Then, waggle runs are clustered following a *HAC* approach, similar to the *AM* (see S2 Fig). In this case, the clustering process is carried out in a three-dimensional data space defined by the axes *X* and *Y* of the comb surface and a third axis *T* of time of occurrence. This way, each *WR* can be represented in the data space by (x, y, t) coordinates based on its comb location and time of occurrence (see S3A Fig). To maintain coherence between spatial and temporal values, the time of occurrence is represented in one fourth of the seconds relative to the beginning of the day.

A threshold Euclidean distance d_{max3} is defined as a parameter for the clustering process (see S3B Fig). The value of the threshold is based on the average drift between *WRs* and the average time gap between consecutive *WRs* (we used the data provided in [6]). We only consider clusters with a minimum of 4 waggle runs as actual dances [23]. Then, we use random sample consensus (*RANSAC*) [48] to find outliers in the distribution of waggle run orientations. Waggle run duration and orientation are then averaged for all inliers and translated to field locations. The mean waggle run duration is translated to meters using a conversion factor, and its orientation is translated to the field with reference to the azimuth at the time of the dance. The duration-to-distance conversion factor was empirically determined by averaging the durations of waggle runs advertising a known feeder (see Discussion for further details).

Experimental validation and results

To evaluate the distance decoding accuracy, we ran the AM on a set of video sequences containing a total of 200 WRs. These videos were recorded for a another research question using different hardware (for details refer to [6]). The duration exported by the AM for each WR was compared to manually labeled ground truth. We found that the AM overestimated WR durations on average by 98 ms, with an SD of 139 ms.

To evaluate the performance of the *OM*, we reviewed the video snippets exported and filtered by the attention module AM and the filter network FN, respectively. Eight coworkers defined the correct waggle run orientation for a set of 200 waggle runs. A custom user interface allowed tracing a line that best fits the dancer's body (see S2 Text). The reference angle for each waggle run was defined as the average of the eight manually extracted angles. The *OM* performed with an average error of -2.92˚ and a SD of 7.37˚, close to the SD of 6.66˚ observed in the human-generated data (further details in S3 Text).
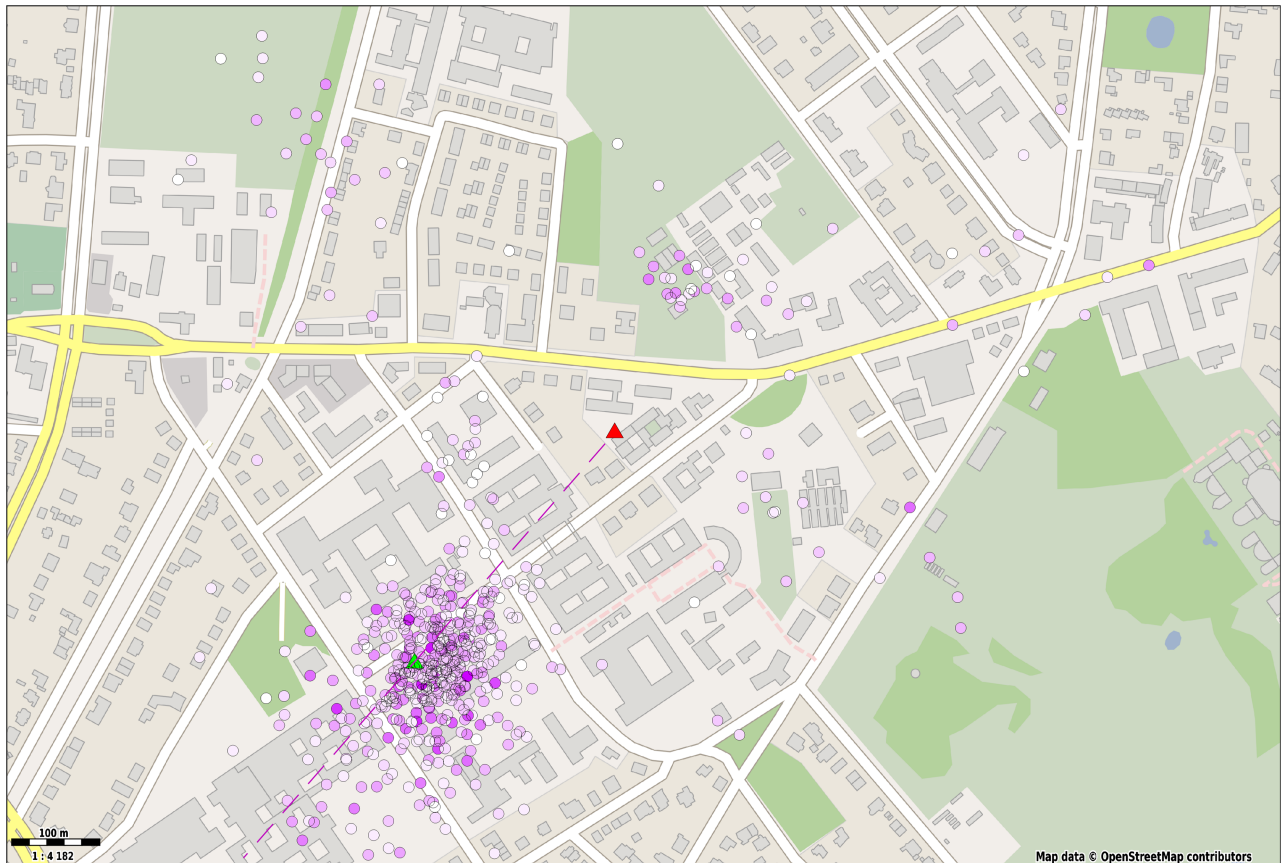
Fig 6. Detected dances mapped back to the field. The average duration and orientation of four or more waggle runs per dance were translated to a field coordinate. The spots' color saturation denotes the number of waggle runs in the dance, white corresponding to the dances with four waggle runs and deep purple to the those with the maximum number (17 for this particular data set). A linear mapping was used to convert waggle duration to distance from the hive. The hive and feeder positions are depicted with a red and green triangle, respectively. The dashed line represents the average direction of all dances (Map data copyrighted OpenStreetMap contributors and available from https://www.openstreetmap.org [49]).

https://doi.org/10.1371/journal.pone.0188626.g006

To illustrate the use cases of the automatic decoding of waggle dances we mapped all dances detected by our system during a period of 5 hours. The data was collected from a honey bee colony kept under constant observation during the summer of 2016. A group of foragers from the colony was trained to an artificial feeder placed 342 m southwest of the hive. Fig 6 depicts the distribution of coordinates converted from 571 dances the system detected. The color saturation of each circle encodes the number of waggle runs associated (5.8 WR on average). Since our bees were allowed to forage from other food sources, not all of the detected dances point towards the artificial feeder. However, most of the detected dances cluster around the feeding site. By averaging the direction of all 571 dances we obtained a very precise match with the artificial feeder's direction, with an angular error below 2.35. If we select only dances in an interval of ±45° around the feeder direction we obtain an angular deviation of 2.33° ± 11.12° on the dance level and 2.68° ± 14.37° on the waggle level.

Discussion

We have presented the first automatic waggle dance detection and decoding system. It is open source and available for free. It does not require expensive camera hardware and works with

standard desktop computers. The system can be used continuously for months, and its accuracy (μ: -3.3˚ ± 5.5˚) is close to human performance (μ: 0˚ ± 3.7˚).

We investigated the possible error sources and visually inspected waggle runs decoded with large error. Before the outlier detection step we find several waggle orientations decoded with an 180˚ error. The orientation of 9% of the waggle runs in our test set were incorrectly flipped. This flip error is corrected in the mapping module by clustering waggle runs to dances and removing angular outliers with *RANSAC*. Once the outliers have been removed, our system performs with an average error of -2.0˚ ± 6.1˚ on the waggle run level. In this process, we discard dances with no strong mode in its waggle orientation distribution and, theoretically, it is therefore possible that a few undiscarded dances contain only flipped waggle runs.

In most of these examples the dancing bee was partly occluded or wagged her body for only short durations, i.e. there is almost no forward motion visible. The forward motion is the central feature in the orientation reader module used to disambiguate the direction. Recognizing anatomical features of the dancing bee, such as the head or abdomen, could help reducing this common error. With falling costs for better camera and computing hardware in mind, we, however, think that, although some bees just don't move forward while wagging, using a higher spatial resolution will likely resolve most of the detection and decoding errors we have described.

The error of the distance decoding could only be assessed for the offline mode of operation, i.e. on prerecorded videos. We found a systematic error that can be ignored with a properly calibrated system. The standard deviation of the waggle duration error was found to be 139 ms. This result highly depends on the choice of the threshold value t_h. We determined the default value of t_h empirically for optimal waggle detection accuracy. We did not explicitly optimize this parameter for more accurate distance decoding.

Bees encode accumulated optical flow rather than metric distances in their waggle runs. Neither the internal calibration, nor the external factors that influence a bee's perception of optical flow were assessable. Hence we calibrated our system with the collective calibration of the very colony under observation: We extracted all waggle runs signaling the location of our artificial feeder (±10˚) and averaged all waggle durations (μ = 582.79 ms ± 196.10 ms). It is unlikely that the set of waggle runs could have contained waggle runs signalling other feeders since natural food sources were scarce in that time of year. This notion is supported by a coefficient of variation of ≈ 0.34, consistent with the value observed by Landgraf et al. in [6].

Given the high accuracy of the method, why do the projected dances in Fig 6 exhibit such a large spread? We inspected random samples and found that not all of the dances advertised our feeding site. Thus, the vector endpoint distribution shows smaller clusters that likely represent natural food sources. The variation of the dance points around the feeding site is correctly reproduced with a large part of the variation originating from the animals themselves. This imprecision is well-know and caused by waggle runs missing the correct direction, with alternating sign. The difference between consecutive waggle runs, or divergence, is surprisingly large and has been studied previously [4, 50, 51]. The divergence correlates negatively with distance to the advertised goal, i.e. it is largest for short distances. In a previous work [6], we analyzed dances to a 215 m distant food source and tracked the motion of all dancers, corrected the tracker manually whenever necessary, and computed the distribution of waggle directions of over 1000 waggle runs. We found that although the average waggle orientation was surprisingly accurate, it was astonishingly imprecise ($\mu = -0.03$˚, $\sigma = 28.06$˚). A similar result was obtained in the present study. The average direction of all dances matches the direction to the feeding station closely ($\Delta = 2.68$˚) with a standard deviation of $\sigma = 14.37$. The spread of dance endpoints, however, is smaller due to the integration of at least four waggle runs ($\sigma = 11.12$˚. Using single waggles or short dances to pinpoint foraging locations of individuals can

therefore unlikely be accurate and it is clear that the number of dances to be mapped needs to be tuned to the given scientific context and environmental structure. We excluded waggle detections shorter than 200 ms, a timespan that would contain less than three body oscillations. Remarkably, bees shake their body in short pulses quite frequently even in non-dance behaviors and hence, the number of false positive detections increases with lower thresholds. Round dances, a dance type performed to advertise nearby resources [1], may also contain short waggle portions. Although our system may be able to detect these, the waggle oscillation is an unreliable feature for round dances. We therefore explicitly focus on waggle dances. Note that the sharp cutoff of dance detections close to the hive in Fig 6 stems from discarding short waggle runs.

The presented system is unique in its approach and capabilities. There are, however, still some features missing that might be added in the future. The mapping module, e.g., does not yet extract and visualize the profitability of a food source. One could, e.g. calculate the return run duration in the clustering step and use a color coding scheme to encode this information into the map. Bee dances also exhibit a systematic angular error that depends on the waggle orientation on the comb ("Restmissweisung", [1]). To improve mapping accuracy, we plan to add a correction step to the mapping module.

The proposed system consists of multiple modules executable as command line programs. Although well documented, this might seem impractical or even obfuscating to the end-user. We are thus developing a graphical user interface to be published in the near future. We currently investigate whether a deep convolutional network is able to extract the relevant image features. If successful, this would enable us to merge the filter network module and the orientation reader, therefore reducing the system's complexity for the user. For the future, we envision an entirely neural system for all the described stages. We also think the solution could be ported to mobile devices. This would enable users an easier setup. Dance orientations could be corrected by reading the direction of gravity directly from the built-in accelerometer. We would like to encourage biologists to use our system and report issues that they face in experiments. Interested software developers are invited to help improving existing features or implementing new ones.

Supporting information

S1 Text. Specifications of the recording setup used during reported experiments. This document contains further information on the recording setup, with an emphasis on technical details.
(PDF)

S2 Text. Further details on the software modules. This document contains diagrams and detailed information on the functioning of the software modules.
(PDF)

S3 Text. Error distributions at multiple levels of analysis. This document provides additional results supporting the case of study presented in the experimental validation and results section.
(PDF)

S1 Fig. A convolutional neural network was used to filter the detections of the AM. The raw sequences of images are processed by two stacked 3D convolution layers with SELU nonlinearities. The outputs of the second convolutional layer are flattened using average pooling on all three dimensions. A final fully connected layer with a sigmoid nonlinearity computes the probability of the sequence being a dance or not. Dropout is applied after the average

pooling operation to reduce overfitting.
(PNG)

S2 Fig. Dendrogram representing the hierarchical agglomerative clustering of 200 waggle runs. The dendrogram is a graphic representation of the clustering process. Each observation starts it its own cluster, at each iteration the two clusters closer to each other are merged, this process is performed recursively till only one cluster remains. We set a threshold distance for clusters to be merged, all clusters generated to the point this threshold is reached are regarded as dances and their constituent elements as their waggle runs.
(PNG)

S3 Fig. Representation of *WRs* in the data space XYT. (A) Representation of a set of 200 *WRs* in the XYT data space, where values in the axes *X* and *Y* are defined by their comb location, and in axis *T* by their time of occurrence. (B) *WRs* within a maximum Euclidean distance of d_{max3} are clustered together and regarded as dances.
(PNG)

S1 Movie. Automatic detection and decoding of honey bee waggle dance: Examples of correct and false detections. A collection of video snippets that were detected by the Attention Module as waggle runs.
(WMV)

Acknowledgments

Author Contributions

Conceptualization: Fernando Wario, Tim Landgraf.

Data curation: Fernando Wario, Benjamin Wild, Tim Landgraf.

Formal analysis: Fernando Wario.

Methodology: Fernando Wario, Tim Landgraf.

Project administration: Raúl Rojas, Tim Landgraf.

Resources: Raúl Rojas, Tim Landgraf.

Software: Fernando Wario, Benjamin Wild, Tim Landgraf.

Supervision: Raúl Rojas, Tim Landgraf.

Validation: Fernando Wario.

Writing – original draft: Fernando Wario, Tim Landgraf.

Writing – review & editing: Fernando Wario, Benjamin Wild, Raúl Rojas, Tim Landgraf.

References

1. von Frisch K. Tanzsprache und Orientierung der Bienen. Berlin: Springer; 1965. Available from: http://www.amazon.de/Tanzsprache-Orientierung-Bienen-Karl-Frisch/dp/3642949177.

2. Seeley TD. The Wisdom of the Hive. Harvard University Press; 1995.

3. Grüter C, Balbuena MS, Farina WM. Informational conflicts created by the waggle dance. Proceedings Biological sciences / The Royal Society. 2008; 275(March):1321–1327.

4. Weidenmüller A, Seeley TD. Imprecision in waggle dances of the honeybee (Apis mellifera) for nearby food sources: Error or adaptation? Behavioral Ecology and Sociobiology. 1999; 46(3):190–199.

5. Tanner DA, Visscher PK. Adaptation or constraint? Reference-dependent scatter in honey bee dances. Behavioral Ecology and Sociobiology. 2010; 64(7):1081–1086. https://doi.org/10.1007/s00265-010-0922-3 PMID: 20585382

6. Landgraf T, Rojas R, Nguyen H, Kriegel F, Stettin K. Analysis of the waggle dance motion of honeybees for the design of a biomimetic honeybee robot. PloS one. 2011; 6(8):e21354. https://doi.org/10.1371/journal.pone.0021354 PMID: 21857906

7. Esch HE, Burns JE. Distance estimation by foraging honeybees. The Journal of experimental biology. 1996; 199(Pt 1):155–62. PMID: 9317542

8. Esch HE, Zhang S, Srinivasan MV, Tautz J. Honeybee dances communicate distances measured by optic flow. Nature. 2001; 411(6837):581–583. https://doi.org/10.1038/35079072 PMID: 11385571

9. Dacke M, Srinivasan MV. Two odometers in honeybees? Journal of Experimental Biology. 2008; 211 (20). https://doi.org/10.1242/jeb.021022 PMID: 18840662

10. Seeley TD, Mikheyev AS, Pagano GJ. Dancing bees tune both duration and rate of waggle-run production in relation to nectar-source profitability. Journal of Comparative Physiology - A Sensory, Neural, and Behavioral Physiology. 2000; 186(9):813–819. https://doi.org/10.1007/s003590000134

11. Biesmeijer JC, Seeley TD. The use of waggle dance information by honey bees throughout their foraging careers. Behavioral Ecology and Sociobiology. 2005; 59:133–142.

12. Riley JR, Greggers U, Smith aD, Reynolds DR, Menzel R. The flight paths of honeybees recruited by the waggle dance. Nature. 2005; 435(7039):205–207. https://doi.org/10.1038/nature03526 PMID: 15889092

13. Menzel R, Kirbach A, Haass WD, Fischer B, Fuchs J, Koblofsky M, et al. A common frame of reference for learned and communicated vectors in honeybee navigation. Current Biology. 2011; 21(8):645–650. https://doi.org/10.1016/j.cub.2011.02.039 PMID: 21474313

14. Al Toufailia H, Couvillon MJ, Ratnieks FLW, Grüter C. Honey bee waggle dance communication: Signal meaning and signal noise affect dance follower behaviour. Behavioral Ecology and Sociobiology. 2013; 67(4):549–556. https://doi.org/10.1007/s00265-012-1474-5

15. von Frisch K. Die Tänze der Bienen. Österreich Zool Z. 1946; 1:1–48.

16. Visscher PK, Seeley TD. Foraging Strategy of Honeybee Colonies in a Temperate Deciduous Forest. Ecology. 1982; 63(6)(6):1790–1801. https://doi.org/10.2307/1940121

17. Barron A, Srinivasan MV. Visual regulation of ground speed and headwind compensation in freely flying honey bees (Apis mellifera L.). Journal of Experimental Biology. 2006; 209(5):978–984. https://doi.org/10.1242/jeb.02085 PMID: 16481586

18. Couvillon MJ, Schürch R, Ratnieks FLW. Waggle Dance Distances as Integrative Indicators of Seasonal Foraging Challenges. PLoS ONE. 2014; 9(4):e93495. https://doi.org/10.1371/journal.pone.0093495 PMID: 24695678

19. Balfour NJ, Ratnieks FLW. Using the waggle dance to determine the spatial ecology of honey bees during commercial crop pollination. Agricultural and Forest Entomology. 2016;

20. Haldane JBS, Spurway H. A statistical analysis of communication in"Apis mellifera"ï¿½ and a comparison with communication in other animals. Insectes Sociaux. 1954; 1(3):247–283. https://doi.org/10.1007/BF02222949

21. Schürch R, Ratnieks FLW. The spatial information content of the honey bee waggle dance. Frontiers in Ecology and Evolution. 2015; 3:22.

22. De Marco RJ, Gurevitz JM, Menzel R. Variability in the encoding of spatial information by dancing bees. The Journal of experimental biology. 2008; 211(Pt 10):1635–1644. https://doi.org/10.1242/jeb.013425 PMID: 18456891

23. Couvillon MJ, Riddell Pearce FC, Harris-Jones EL, Kuepfer AM, Mackenzie-Smith SJ, Rozario LA, et al. Intra-dance variation among waggle runs and the design of efficient protocols for honey bee dance decoding. Biology open. 2012; 1(5):467–72. https://doi.org/10.1242/bio.20121099 PMID: 23213438

24. Landgraf T, Oertel M, Kirbach A, Menzel R, Rojas R. Imitation of the honeybee dance communication system by means of a biomimetic robot. In: Lecture Notes in Computer Science. vol. 7375 LNAI. Springer-Verlag; 2012. p. 132–143.

25. Seeley TD. Social foraging by honeybees: how colonies allocate foragers among patches of flowers. Behavioral Ecology and Sociobiology. 1986; 19(5):343–354. https://doi.org/10.1007/BF00295707

26. Camazine S, Sneyd J. A model of collective nectar source selection by honey bees: Self-organization through simple rules. Journal of Theoretical Biology. 1991; 149(4):547–571. https://doi.org/10.1016/S0022-5193(05)80098-0

27. de Vries H, Biesmeijer JC. Modelling collective foraging by means of individual behaviour rules in honey-bees. Behavioral Ecology and Sociobiology. 1998; 44(2):109–124. https://doi.org/10.1007/s002650050522

28. Seeley TD, Buhrman SC. Nest-site selection in honey bees: How well do swarms implement the "best-of-N" decision rule? Behavioral Ecology and Sociobiology. 2001; 49:416–427. https://doi.org/10.1007/s002650000299

29. Passino KM, Seeley TD. Modeling and analysis of nest-site selection by honeybee swarms: the speed and accuracy trade-off. Behav Ecol Sociobiol. 2006; 59:427–442. https://doi.org/10.1007/s00265-005-0067-y

30. Reina A, Valentini G, Fernández-Oto C, Dorigo M, Trianni V. A Design Pattern for Decentralised Decision Making. PLOS ONE. 2015; 10(10):e0140950. https://doi.org/10.1371/journal.pone.0140950 PMID: 26496359

31. Wario F, Wild B, Couvillon MJ, Rojas R, Landgraf T. Automatic methods for long-term tracking and the detection and decoding of communication dances in honeybees. Frontiers in Ecology and Evolution. 2015; 3(September):1–14.

32. Khan Z, Balch T, Dellaert F. A Rao-Blackwellized Particle Filter for Eigentracking. In: 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. vol. 2. IEEE; 2004. p. 980–986. Available from: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1315271.

33. Kimura T, Ohashi M, Okada R, Ikeno H. A new approach for the simultaneous tracking of multiple Honeybees for analysis of hive behavior; 2011. Available from: http://www.scopus.com/inward/record.url?eid=2-s2.0-80054706539&partnerID=40&md5=033421f58e3b80cbb6293f9b98431a29.

34. Landgraf T, Rojas R. Tracking honey bee dances from sparse optical flow fields. Free University Berlin; 2007. June.

35. Feldman A, Balch T. Automatic Identification of Bee Movement. Georgia: Georgia Institute of Technology; 2003. Available from: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.122.300&rep=rep1&type=pdf.

36. Kabra M, Robie AA, Rivera-Alba M, Branson S, Branson K. JAABA: interactive machine learning for automatic annotation of animal behavior. Nature Methods. 2012; 10(1):64–67. https://doi.org/10.1038/nmeth.2281 PMID: 23202433

37. Oh SM, Rehg JM, Balch T, Dellaert F. Learning and Inferring Motion Patterns using Parametric Segmental Switching Linear Dynamic Systems. International Journal of Computer Vision. 2008; 77(1-3):103–124. https://doi.org/10.1007/s11263-007-0062-z

38. Feldman A, Balch T. Representing Honey Bee Behavior for Recognition Using Human Trainable Models. Adaptive Behavior. 2004; 12(3-4):241–250. https://doi.org/10.1177/105971230401200309

39. Gil M, De Marco RJ. Decoding information in the honeybee dance: revisiting the tactile hypothesis. Animal Behaviour. 2010; 80:887–894. https://doi.org/10.1016/j.anbehav.2010.08.012

40. Cooley JW, Tukey JW. An Algorithm for the Machine Calculation of Complex Fourier Series. Mathematics of Computation. 1965; 19(90):297. https://doi.org/10.1090/S0025-5718-1965-0178586-1

41. Sibson R. SLINK: An optimally efficient algorithm for the single-link cluster method. The Computer Journal. 1973; 16(1):30–34. https://doi.org/10.1093/comjnl/16.1.30

42. Ji S, Xu W, Yang M, Yu K. 3D Convolutional Neural Networks for Human Action Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2013; 35(1):221–231. https://doi.org/10.1109/TPAMI.2012.59 PMID: 22392705

43. Tran D, Bourdev L, Fergus R, Torresani L, Paluri M. Learning Spatiotemporal Features with 3D Convolutional Networks. arXiv preprint. 2014;

44. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. Journal of Machine Learning Research. 2014; 15:1929–1958.

45. Klambauer G, Unterthiner T, Mayr A, Hochreiter S. Self-Normalizing Neural Networks. arXiv preprint. 2017;

46. Kingma DP, Ba J. Adam: A Method for Stochastic Optimization. In: Proceedings of the 3rd International Conference on Learning Representations (ICLR); 2014. Available from: http://arxiv.org/abs/1412.6980.

47. Bracewell R. The Projection-Slice Theorem. In: Fourier Analysis and Imaging. Boston, MA: Springer US; 2003. p. 493–504. Available from: http://link.springer.com/10.1007/978-1-4419-8963-5_14.

48. Fischler MA, Bolles RC. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM. 1981; 24(6):381–395. https://doi.org/10.1145/358669.358692

49. OpenStreetMap contributors. Planet dump retrieved from https://planet.osm.org; 2017. https://www.openstreetmap.org.

50. Towne WF, Gould JL. The spatial precision of the honey bees' dance communication. Journal of Insect Behavior. 1988; 1(2):129–155. https://doi.org/10.1007/BF01052234

51. Tanner DA, Visscher K. Do honey bees tune error in their dances in nectar-foraging and house-hunting? Behavioral Ecology and Sociobiology. 2006; 59(4):571–576. https://doi.org/10.1007/s00265-005-0082-z