Description and validation of the LocoWhisk system: Quantifying rodent exploratory, sensory and motor behaviours

David Gillespie[a], Moi Hoon Yap[b], Brett M. Hewitt[b,c,d], Heather Driscoll[a,e], Ugne Simanaviciute[c,f], Emma F. Hodson-Tole[g], Robyn A. Grant[c*]

a School of Engineering, Manchester Metropolitan University, Manchester, UK

b School of Computing, Mathematics and Digital Technology, Manchester Metropolitan University, Manchester, UK

c School of Science and Environment, Manchester Metropolitan University, Manchester, UK

d CSols Ltd, Runcorn, Cheshire, WA7 4QX, UK

e Advanced Manufacturing Research Center, University of Sheffield, Sheffield, UK

f School of Biological Sciences, University of Manchester, Manchester, UK

g Musculoskeletal Science and Sports Medicine Research Centre, Dpt. Lifesciences, Manchester Metropolitan University, Manchester, UK

* Corresponding Author: Robyn.A.Grant: robyn.grant@mmu.ac.uk

Abstract

Background: Previous studies have demonstrated that analysing whisker movements and locomotion allows us to quantify the behavioural consequences of sensory, motor and cognitive deficits in rodents. Independent whisker and feet trackers exist but there is no fully-automated, open-source software and hardware solution, that measures both whisker movements and gait.

New Method: We present the LocoWhisk arena and new accompanying software (ARTv2) that allows the automatic detection and measurement of both whisker and gait information from high-speed video footage.

Results: We demonstrate the new whisker and foot detector algorithms on high-speed video footage of freely moving small mammals, and show that whisker movement and gait measurements collected in the LocoWhisk arena are similar to previously reported values in the literature.

Comparison with existing method(s): We demonstrate that the whisker and foot detector algorithms, are comparable in accuracy, and in some cases significantly better, than readily available software and manual trackers.

Conclusion: The LocoWhisk system enables the collection of quantitative data from whisker movements and locomotion in freely behaving rodents. The software automatically records both whisker and gait information and provides added statistical tools to analyse the data. We hope the LocoWhisk system and software will serve as a solid foundation from which to support future research in whisker and gait analysis.

Keywords: locomotion, whisking, automated tracking, software development, image processing

Introduction

Many small laboratory mammals (rats, mice, guinea pigs, shrews, opossums) rely on their whiskers as their primary sense to guide locomotion and environmental exploration (Arkley et al., 2014; Mitchinson et al., 2011; Grant et al., 2009; Grant et al., 2018a). While the whisker system is an established model for cortical sensory pattern formation and information processing, recent advances in measuring whisker behaviours have enabled it to be suggested as a candidate for also measuring motor control (Grant et al., 2014; Garland et al., 2018). Aspects of whisker movements can also be reliably affected during a variety of sensory, motor and cognitive disorders (Simanaviciute et al., in this issue), including rodent models of Amyotrophic Lateral Sclerosis (Grant et al., 2014), anxiety (Grant et al., 2016), Huntington's disease (Garland et al., 2018) and Alzheimer's disease (Grant et al., 2018b). In the case of Huntington's disease and Alzheimer's disease, mice revealed whisker movement deficits earlier than any other motor behaviour test, including gait assessments (Garland et al., 2018). This demonstrates the utility of measuring whisker movements to quantify motor and cognitive declines.

Freely available computer vision libraries (such as OpenCV (Bradski and Kaehler, 2008)) have allowed for the development of several automated rodent whisker trackers, which measure whisker movements from high-speed video. These include: i) WhiskerTracker, which requires infrared-lighting to identify eye positions (Knutsen et al., 2005); ii) an unsupervised tracker (Voigts et al., 2008), which can track a single row of whiskers when all the other whiskers have been trimmed; iii) Whisk, which has been designed for a head-fixed set-up (O'Connor et al., 2010; Clack et al., 2012); iv) Whiskerman (Bale et al., 2015) which tracks and extracts parameters from a single whisker automatically where other whiskers are present and not trimmed; and v) the BIOTACT Whisker Tracking Tool (BWTT) (Perkon et al., 2011), which is only semi-automated, and settings cannot be readjusted through the video to account for areas of high or low brightness and contrast. These whisker trackers have been created by researchers and made freely available (Clack et al., 2012; Knutsen et al., 2005; Perkon et al., 2011); however, there does not yet exist a robust, fully-automated whisker tracker, that can detect whiskers within an intact whisker field in freely moving animals.

Foot tracking software exists that tracks and records gait information for rodents, including the Catwalk XT (Noldus Information Technology BV, Wageningen, The Netherlands), MouseWalker (Mendes et al., 2015) and AGATHA (Kloefkorn et al., 2017). However, these systems often require two views (below and side-on), very specific and expensive experimental set-ups (such as the Catwalk XT) and the feet to be clearly imaged, preferably in colour. These software packages rarely work with videos that have been imaged to be used for whisker tracking (i.e. low contrast, monochrome videos). Therefore, it is not possible to conduct whisker and foot tracking together. Used in tandem with locomotion (an established behavioural measure), whisker movements could give an added dimension to quantifying motor control, since these are closely coupled during development (Grant et al., 2012). In addition, it has also been demonstrated that whisker movements can be linked to an animal's locomotion speed, and adjusted according to their environmental experience and attention (Arkley et al., 2014). Therefore, locomotion and whisker movements are closely linked and should be studied in tandem.

This has prompted the development of the LocoWhisk system, that can simultaneously measure and quantify rodent whisker and locomotion movements. The primary ambition of the LocoWhisk system is to provide a new, non-invasive assessment of animal health and welfare, through the provision of open-source software and hardware to quantify rodent whisker and locomotion

movements. In this article, we introduce new methods for automatically detecting whiskers and feet in small mammals for use with the LocoWhisk arena, and also validate these detection algorithms. For the whisker detector we evaluate its performance against state-of-the-art software, and for the foot detection algorithm it is validated against the ground truth that we produced from manual annotations.

Methods

LocoWhisk Arena

The LocoWhisk arena is a bespoke, low cost hardware system that enables monochrome high-speed video footage to capture the whiskers and feet simultaneously. The LocoWhisk arena consisted of an aluminium support frame, made of Bosch Rexroth compatible profiles (Figure 1a). These profiles enclosed a glass plate floor, and supported a Perspex arena. The glass plate (6 mm thick) was edged with a strip of 12 V red LEDs (60-100 lumen per metre at 625 nm) around its perimeter, creating a pedobarograph. Clear float glass was selected for the pedobarograph for its high transmission index (Corr et al., 1998). Red LEDs were used as rodents are less sensitive to red light (Dauchy et al., 2015; Peirson et al., 2018). The glass pedobarograph floor was enclosed by sections of Bosch Rexroth compatible profiles, connected to the support frame with a hinge. The floor could be inclined up to 30 degrees (Figure 1a), although all recordings were made on a flat floor in this study (at zero degrees).

Animals were placed into a Perspex arena (300x500x150 mm for rats, and 300x200x120 mm for the mice and other species). The arena had a hinged lid, and open bottom, so animals directly contacted the pedobarograph floor with their feet. Contact between the animals' feet and the glass plate disrupted the total internal reflection of the red light. When a foot contact was made with the pedobarograph, the contacting area lit up and could be viewed beneath the feet when filmed from below. The intensity of the light is proportional to the pressure applied. When no foot contact was made with the surface of the pedobarograph, the red lights internally reflected through the glass, and as such could not be seen when viewed from below. Pedobarographs often require a layer of textured paper for surface contact to cause a change in the refractive index of the glass; this was not required for the animals as the surface profile of their feet was sufficient to give rise to the pedobarograph effect. This was of benefit for the LocoWhisk arena as it enabled the whiskers to still be visible while simultaneously imaging the feet.

To image the whiskers, a high-speed, high-resolution video camera (Phantom Miro ex2) was positioned below the floor, filming at 500 fps at 640x480 pixels. An infrared light slate (PHLOX SLLUB Backlight 850 nm) was placed on top of the arena to illuminate the whiskers. Multiple video clips were collected opportunistically (by manual trigger) when each animal passed in to the view of the camera, and ranged from 0.6 to 1.6 s in length. Recording stopped when the camera memory was full or the animal stopped exploring. All videos were saved in Audio Video Interleave (avi) format and uploaded to the LocoWhisk ARTv2 software, which automatically selected trackable videos and frames. Trackable videos and frames were identified as when the mouse's head, was clearly shown in the image, and when the mouse was not rearing, which obscured the nose tip (as per Hewitt et al., 2016; Grant et al. 2014).

To validate and measure the accuracy of the whisker detector algorithm, video footage was used that had been previously recorded in Grant et al. (2013) and Grant et al. (2018a). Video clips were randomly selected for tracking and included 9 clips of brown rats (*Rattus norvegicus*), 36 clips of house mice (*Mus musculus*), 1 clip of harvest mouse (*Micromys minutus*), and 4 clips of a gray short

tailed opossum (*Monodelphis domestica*). The opossum was kept and filmed at the university of Trieste, Italy, and all other species were filmed at the Wildwood Trust in Kent, UK. All animals were adult, and a mixture of males and females were used, animals were assumed to be sexually monomorphic.

To trial the arena, whisker outputs and foot detector algorithm, 4 adult brown rats (*Rattus norvegicus*) and 2 adult wood mice (*Apodemus sylvaticus*) were tested at the Wildwood Trust, Kent. 4 video clips were analysed from rat and 3 from mice for the validation of the foot detector and software out*p*uts. All procedures were approved by the local ethics committees at the Wildwood Trust, University of Sheffield, University of Trieste and Manchester Metropolitan University, in accordance with regulations issued by the UK Home Office.

Software

The software associated with the LocoWhisk arena, has been added to the existing ART software (Hewitt et al., 2018) and will be referred to in this article as ARTv2. ART is an Automated Rodent Tracker capable of generating qualitative and quantitative data of rodent movement and exploration, as well as tracking the distance and speed of the rodent movements (Hewitt et al. 2018). ARTv2 is an expansion of this software, building upon the qualitative and quantitative information generated in ART to incorporate whisker movements and gait measurement, as the animal moves within the LocoWhisk arena. It follows the same ethos as ART, being open-source and able to provide similar or better accuracy for whisker tracking than commercial software with minimal user interactions. It is also possible to use previous videos processed by ART to attempt to detect whiskers and foot information within the clip.

Whisker Detection Algorithm

The whisker detection algorithm incorporated into ARTv2, combined several image-processing techniques to enhance the whiskers and separate them from background noise. As whiskers can be difficult to detect against the background of the LocoWhisk arena, a region of interest (ROI) was extracted from the image with a minimal amount of other objects or artefacts in the scene. The ROI was defined as the region between 5% to 20% of the resolution outwards of the animal contour (Figure 2a). This was to avoid noise generated from the immediate surroundings around the animal, and to cover a significant amount of whiskers. A circular segment was added to exclude the nosetip from the binary mask. This mask was then used to extract the ROI from the original image as shown in Figure 2b.

With all the background objects removed from the scene, the ROI was then inverted, and pre-processed by linear normalisation (Rosenfeld, 1976) as in Equation 1, where ROI is denoted as *I*, *newMin* is the minimum pixel value and *newMax* is the maximum of the pixel value detected within the image *I*. Pixel value is the greyscale pixel intensity value, and *Min* and *Max* are defined as the minimum and maximum pixel value of an image, which are 0 and 255 respectively for 8-bit images. I_N was then calculated using Equation 2.

$$I: \{X \subseteq R^n\} \rightarrow \{Min \dots Max\} \qquad (1)$$

$$I_N = (I - Min)\frac{newMax - newMin}{Max - Min} + newMin \qquad (2)$$

The equalisation was kept linear as any other form of equalisation introduced too much noise into the image and created additional line-like segments (Figure 2c). A threshold value was calculated as the mean pixel intensity of the image, and every pixel from the source image (defined as *src* in Equation 3) below this value was thresholded to zero (Gonzalez and Woods, 2002) (Equation 3). This process was repeated with an increasing threshold value, until the mean intensity of the masked image fell below 10% of the bit depth (25 intensity, for 8-bit images) (Figure 2d). This handles images of varying brightness, and ensures noisy parts of the image are not picked up as potential whisker locations.

$$dst(x,y) = \begin{cases} src(x,y), & if\ src(x,y) > threshold \\ 0, & otherwise \end{cases} \quad (3)$$

After the image was thresholded with equation 3, the whiskers were enhanced, but also shown as several pixels wide (Figure 2d), with some noise still present. In order to highlight the whiskers further, probable whisker locations were located. This was done by repeating the process as above, but instead of using a mask constructed of an inner and outer extension, a mask of only 1 pixel in width was constructed for all values between the inner and outer extension distances. For each mask, a Gaussian filter was applied to remove noise and to create a "smoother" gradient profile of the pixels. Non-maxima suppression (Devernay, 1995) was then applied to the entire kernel, rather than in a specified direction, to leave only the largest intensity pixel. The size of the kernel was a trade-off between sensitivity and specificity. Kernel size was determined as 2% of the resolution. The resulting pixels can be considered a probable location for a whisker position. The process was repeated several times with the contour being extended by a further pixel each time, leaving an area several pixels wide with probable whisker locations shown in Figure 3a. A median subset of the extension distances was chosen to look for the final whisker locations.

Each pixel from the non-maxima suppressed image was chosen, and a line mask was created for every angle (0 - 180◦), the length of which was kept short so curved whiskers could still be approximated by a straight line. This mask was then applied to an image consisting of the potential whisker locations (Figure 3a) and added to the original extracted image (Figure 3b). The probable whisker locations increase the intensity considerably, thus giving a much larger mean value when one of the orientations falls in the correct position. If the mean intensity increased above a certain pixel value threshold of grayscale intensity (112 pixel value), then the orientation with the largest response was saved for that pixel. The previous step will have generated several potential whisker locations (Figure 3c), some of which will be for the same whisker, it was therefore necessary to reduce these potential locations to just one for each whisker (Figure 3d). The median extension distance from the subset (15 in this case) was then chosen as the point to determine if the current pixel was actually a whisker and if the orientation was correct. All potential whisker lines within a 7 × 7 area were selected. If the point was a whisker, it can be expected that several lines will have been found in the surrounding area, all with a similar angle. If the number of lines fell below a certain threshold (50% of the extended ranges, 4 in this case), then it was instantly discarded. If there were 4 or more lines, they were ordered by their angle. The first and last 25% were removed to remove any large outliers, then the line with the largest mean intensity value was chosen to represent the whisker, while all other lines were suppressed (Figure 3d). This approach works well in practice, but a limitation may be that whiskers that are extremely close together will only be identified as one whisker. While the detection process achieved a very high true positive rate, it also generated a significant amount of false positives. In order to reduce the false positives, rule-based logic was applied to each detected whisker, such that if the whisker was found to be parallel to the orientation of the rodent's face, it was removed. This approach consistently left true positive whiskers.

As the whisker detection algorithm is a part of ARTv2, it is possible to export various statistics and measurements about the animal in the clip. The outputs are documented in Table 1, and include whisking frequency, whisker angle, protraction and retraction speed, spread and amplitude of the whiskers. The mean whisker angular position is one of the most important metrics calculated in ARTv2 that the other outputs rely on. It is defined in this paper as the angle between the centre line of the rodents head and the tangent to a given point on a whisker shaft, as shown in Figure 4 and is used in multiple studies (Perkon et al., 2011; Clack et al., 2012; Pammer et al., 2013; Ahissar and Knutsen, 2008; Sofroniew et al., 2014).

Foot detection algorithm

The foot detection algorithm combined several image-processing techniques to detect rodent feet, label the front and hind feet on either side of the body, and also persistently track each identified foot through a video. This was accomplished in two phases, one to detect the feet, while the second phase automatically labelled and tracked the correct foot.

Due to rodents being very small and light, and priority given to imaging the whiskers, it can be difficult to detect foot placements from the LocoWhisk arena pedobarograph. In order to clearly highlight the feet of the rodent, the input image was converted to log colour space, as can be seen in Figure 5. This image, in combination with the animals body contour (which is processed as part of ARTv2 and detailed in Hewitt et al. (2018), was used to create a mask which was eroded using OpenCV's erode function (Bradski and Kaehler, 2008). This erosion function removes visual artefacts so that only the feet can be seen, as displayed (Figure 5b.). Any foot that lies outside the body contour is not recorded. The resolution of the mask was then reduced to remove visual artefacts and allow for foot contours to be more easily detected. The mask was processed by the OpenCV findContour (Bradski and Kaehler, 2008) function, to provide a bounding box and a centre point for each foot (Figure 5d).

To identify which foot is which, two points were created to split the animal into a left and right side. This is accomplished by using the central point of the body contour and an angle to create the head point and a point between the tail and centroid point (Figure 5, e.). The angle is taken from the angle between the centre point position and the nose tip recorded with ARTv2 (Hewitt et al., 2018) using the inverse arctangent formula. If no nose tip points are available, the angle is calculated by taking the angle between the centroid point in the current frame and a centroid point captured in a subsequent frame using the inverse arctangent formula.

With the created points, it was then possible to calculate the perpendicular distance from the cross product between the head point to the centroid position and the foot point to the centroid position. The parallel distance was also calculated from the dot product between these points. If the perpendicular distance was positive, the foot lies on the right side, otherwise it lies on the left side. If the parallel distance was greater than one fifth of the distance to the head point, it is labelled as a front foot, otherwise it is labelled as a hind foot.

To persistently track each foot, every frame of the video is inspected for contours; if there are any they are assigned to either an existing or a new identifier. The identifier contains a unique ID with bounding box and centre position. If there no identifiers are stored within the software, the bounding box is registered and the correct foot identifier is calculated. However, if there are registered identifiers, instead of assigning a new ID, we determine if we can associate the foot contour with one of the existing identifiers. This is accomplished by computing the Euclidean distance between the centre points of the new contours, and existing identifiers centre points. We

associate new contours that have the minimum distance to the existing identifier, and update its values for that frame. However, if there is a contour, that is not close to any of the existing points, it is registered as new foot identifier. To handle identifiers that are no longer discoverable after a certain number of frames, such as when the animal lifts its foot, the identifier is deregistered after a certain amount of frames. This is set as 2 frames in the software (i.e. 4 ms at 500 fps) but can be manually adjusted by the user.

After all of the frames have been processed to detect feet, each individual foot is post processed to remove erroneous data (such as feet recorded to be placed for less than 15 frames). ARTv2, is then able to export the stride length, swing time and stance time for each individual foot as can be seen in table 1. It is also possible with other recorded measurements results to be able to calculate other statistics, such as swing and stance speed, stride distance or pressure placed upon the pedobarograph (Mendes et al., 2015).

LocoWhisk System outputs

Once the whiskers and feet have been identified with visual indicators to allow users to inspect the detected whiskers, feet and body contour, it is possible to export various outputs from the software. The ARTv2 software currently exports 4 output files in comma separated values (.CSV) format, 2 are general statistics and 2 are mean data for whisker and feet detection, respectively. Table 1 contains the outputs and the definitions of how each variable is calculated.

Table 1: The outputs that are calculated by ARTv2 for each clip and on a frame-by-frame basis.

Variable Name	Description
Measures per Clip: Each variable is the mean for each clip and calculated for both left and right side.	
Whisker movement measures:	
Mean Angle (Deg)	The average whisker angular positions from all tracked whiskers.
Frequency (Hz)	An auto-correlogram of the left and right angular positions.
Retraction Speed (Deg/s)	Average retraction velocity (backward moving whiskers) from the angular positions.
Protraction Speed (Deg/s)	Averaged protraction velocity (forward moving whiskers) from the angular positions.
Spread (Deg)	The Standard Deviation of all tracked whiskers.
Amplitude	The Standard Deviation of the angular positions, multiplied by $2x \sqrt{2}$ as suggested in (Grant et al., 2014)
Asymmetry	Left minus right mean whisker angles.
Gait measures:	
Stride Distance(mm)	The mean Euclidean distance between individuals foot placements.
Stance Time (ms)	The amount of time that the foot is placed on the floor.
Swing Time (ms)	The amount of time that the foot is lifted in the air.

Variable Name	Description
Measures per frame:	**Variables that are provided on a frame-by-frame basis for each clip, for the left and right side.**
Whisker and head measures:	
Nose point coordinates	The X,Y coordinates of the nose point.
Mid Head point coordinates	The X,Y coordinates of the midpoint of the head.
Mean whisker Angle	The mean angular position from all tracked whiskers.
Whisker points arrays	An array of the whisker points that were detected in ARTv2.
Gait measures:	
Centroid point of the foot	The centre point of the bounding box for each foot.
Foot Dimensions	The width and height of each foot.
Bounding box of the foot	The coordinates of the bounding box, for each foot is stored for each frame.

Validation of whisker detection

To evaluate the performance of ARTv2's whisker detection algorithm, one image was extracted from the video data (50 in total) of many small mammal species performing locomotion and exploration of static objects placed within the set up. The ability of ARTv2 to detect whiskers was measured by counting how many whiskers were correctly detected (True Positive (*TP*)) in 50 selected frames (images) from the collected videos. To compare ARTv2 to existing trackers, Whisk and BWTT were also run on the same 50 images. The number of whiskers detected by Whisk and BWTT were identified, again, by counting the number of true whiskers detected for each tracker. Manual inspection was used to ensure that the detected whiskers were *TP* detections, and only *TP* detections were reported for each tracker. True negative values were unable to be obtained due to the nature of the images, as this would cover almost the entire ROI and would skew the accuracy results.

Manual annotations (i.e. MWA, Hewitt et al. 2016) and trackers (ARTv2, Whisk, BWTT) all have different standards of whisker identification. They use different methods of identification; this means that they identify different whiskers as well as different lengths and shapes (curves) of each whisker. Therefore, it was not possible to fairly compare the pixel tracking values between the software (i.e. by using an Intersection over the Union (*IoU*) score). Therefore, only the number of accurately *TP* detected whiskers was recorded for this study.

Validation of foot detection

To validate the foot detector in ARTv2, an Intersection over the Union (*IoU*) score was calculated measuring the overlap between predicted foot and ground truth foot annotations. *IoU* is a commonly used metric for object detection (He et al., 2017; Redmon et al., 2016; Lin et al., 2017). Any algorithm that provides predicted coordinates can be evaluated using *IoU*. *IoU* is calculated

using equation 4, which measures the area of overlap between the predicted and ground truth coordinates and dividing by the area encompassed by both a ground truth (manual annotation, or a validated ground truth dataset) and predicted outputs (Real and Vargas, 1996).

$$J(A,B)= (|A \cap B|)/(|A \cup B|)= (|A \cap B|)/(|A|+|B|-|A \cup B|) \quad (4)$$

For foot detection validation, 20 images were selected from our existing database of whisker videos from Brown Rats (*Rattus norvegicus*) (n= 10 images), and wood mice (*Apodemus sylvaticus*) (n=10 images), as these are representative of common laboratory animal feet. Specifically, to evaluate the performance of the foot detector, an *IoU* score was computed using the predicted bounding boxes from the foot detector versus ground truth bounding boxes. The selected images were converted to the Log colour space, and then annotated using the Supervisely tool (Supervise.ly, 2019) to create ground truth bounding boxes for each foot. The bounding boxes, were predicted in the software, using different settings for the mice and rats. The settings used to capture the rat feet predictions included: down sampling the created mask twice, eroding the mask fourteen times with a 3x3 kernel, and excluding any bounding box, with an area less than 60 pixels. As mice feet are a lot smaller and the mice were a lot lighter, the settings were different. The mask was down sampled once, eroded three times with a kernel size of 3x3 to remove visual errors and bounding boxes were excluded with an area less than 10 pixels. Due to the varying nature of the parameters used within the foot-tracking algorithm (image pyramid, erosion, joining contours together, excluding any area outside the animals contour), a complete match to the ground truth boxes is highly unlikely. *IoU* was chosen to evaluate the foot tracking algorithm as it rewards predicted bounding boxes that overlap with ground truth annotations (He et al., 2017). An *IoU* score greater than 0.5 is normally considered a good prediction (He et al., 2017).

Results

Whisker detection results

The mean number of whiskers (*TP*) detected for ARTv2 in 50 images was 15.84 ± 5.9 (n=50)) (Table 2). The mean number of whiskers (*TP*) detected per image for Whisk was 6.8 ± 6.97 (n=50). ARTv2 found significantly more whiskers than Whisk (Wilcoxon Signed Rank: $Z=-6.0927$, $p<0.01$) across the same dataset. The mean number of whiskers detected for BWTT was 14.76 ± 4.7 (n=50).Although the *TP* detection rate for ARTv2 was slightly higher, there was no significant difference between the trackers (Wilcoxon Signed Rank: $Z=-1.0554$, $p=0.2891$).

Although ARTv2 did not significantly improve upon BWTT over the full dataset, its success was largely dependent on the resolution of the whiskers or "zoom level". Some videos showed a larger animal, or an animal that was much closer to the camera, and therefore showed the whiskers with a higher spatial resolution (Z1), whereas others recorded a smaller animal, or an animal that was further away (Z2). ARTv2 performed consistently against all zoom levels, although it performed significantly better on videos with high spatial resolution of the whiskers, than BWTT on the Z1 dataset (Wilcoxon Signed Rank: $Z=-2.6672$, $p<0.001$, $n=13$), whereas there was no significant difference on videos with lower spatial resolution on the Z2 dataset. as shown in Table 2, Figure 6.

Table 2. Results table showing the score (compared to manual tracking), mean number of correct detections, degree of automation, and the number of settings required for each method.

Method	Mean No. of whiskers detected			Automation	Dynamic Settings
	All (n=50)	Z1 (n=13)	Z2 (n=37)		
ARTv2	15.84	22.62	13.5	Fully automated	Calculated per frame
Whisk	6.8	16.69	3.32	Fully automated	Calculated per frame
BWTT	14.76	17.23	13.89	Manual settings at start-up	Calculated on one frame and propagated through clip

Foot detection results

The mean *IoU* score for the foot tracking algorithm over the 20 images was 0.502 ± 0.24 std (n=20). The foot detector itself had a mean *IoU* score of 0.59 (± 0.177 std, n=10) for rats and 0.37 (± 0.261 std, n=10) for mice.

Locowhisk System Outputs

Table 3 contains a summary of the statistics recorded by ARTv2 for the rats and mice for the same clips used during the foot tracking validation (above). Mice moved their whiskers with higher frequency, lower amplitude, and their whiskers were more spread out than rat whiskers (Table 3). As mice tend to be much faster and are smaller than rats, all locomotion measures for mice were lower than those of the rat (Table 3).

Table 3: Mean whisker and gait statistics recorded using ARTv2.

Measures per clip	Rats (n=4) (±std)	Mice (n=3) (±std)
Whisker movement measures:		
Frequency (Hz)	11.04 ± 5.08	19.49 ± 8.02
Mean Angle (Deg)	84.99 ± 13.36	83.74 ± 10.73
Retraction Speed (Deg/ms)	1.274 ± 0.51	1.535 ± 0.79
Protraction Speed (Deg/ms)	1.434 ± 0.68	0.963 ± 0.41
Spread (Deg)	16.21 ± 2.40	20.95 ± 1.01
Amplitude	24.2 ± 9.30	17.36 ± 4.47
Gait measures:		
Fore Feet Stride Distance (mm)	101.5 ± 42.0	56.95 ± 10.3
Hind Feet Stride Distance (mm)	48.52 ± 0.0	51.42 ± 13.93
Fore Feet Stance Time (ms)	302.0 ± 89.75	131.7 ± 34.38
Hind Feet Stance Time(ms)	365.4 ± 248.7	171.3 ± 72.28
Fore Feet Swing Time (ms)	385.2 ± 130.8	498.3 ± 374.0
Hind Feet Swing Time (ms)	476.0 ± 260.3	216.3 ± 40.42

Discussion

This paper presents the LocoWhisk system that enables the non-invasive assessment of rodent exploratory behaviour by measuring rodent whisker and locomotion movements. It provides the methods and validation of the software (ARTv2), and demonstrates that these are robust and accurate when compared to other trackers or manual alternatives.

Whisker detection

The whisker detector algorithm incorporated into ARTv2, performed similarly to BWTT and better than Whisk. Indeed, Artv2 outperformed Whisk on the detection rate for all whiskers, the *TP* rate was significantly higher for the Z2 dataset when compared to Whisk. This highlights ARTv2's potential to automatically detect whiskers which have low spatial resolution. Whisk also had a high false positive rate, with scratches on the arena floor and parts of the animal's fur being detected as whiskers. These false positives may occur because Whisk was designed for use on head restrained rodents with only a single row of whiskers on either side (Clack et al., 2012).

The ARTv2 whisker detector is comparable to BWTT for tracking whiskers in freely moving animals with intact whiskers. It is also fully-automated, whereas BWTT is only semi-automated and required manual intervention for the settings. Although ARTv2 performed as well as BWTT on the Z2 dataset (Table 2), it detected a significantly higher amount of *TP* whiskers on the Z1 dataset. It should also be noted that all the settings in ARTv2 are dynamically computed for each individual frame. BWTT manual settings can have a major impact on whisker detection rates (Figure 8). These settings are then applied across an entire video, meaning if the rodent were to travel from an area of high brightness into an area of low brightness, which is common, the program may begin to fail to detect whiskers. ARTv2 would continue to operate at the same detection rate regardless of brightness variation. We used only one image per video for validation, which mean BWTT was manually set up exactly for that frame, and all other frames are likely to degrade from there.

Foot detection

As well as a fully automated whisker detection, we also present a foot tracking algorithm that is able to correctly label, and detect feet from monochrome footage which is collected to detect whiskers of small rodents (Grant et al., 2018a). The foot detector in ARTv2 was not validated against any other software, due to either the software not being open source (Herold et al., 2016), dependent on two views ((Kloefkorn et al., 2017) or being unable to track feet within the monochrome video (Mendes et al., 2015). However, against manual annotation of still images, it was possible to obtain an *IoU* score of 0.502. It is also possible from the outputs generated from this tracker, to calculate stride length, swing speed, swing time, stance time, for each foot, and even the possibility to present pixel intensity maps for a visual representation on how much pressure the animal is applying to the ground.

The mean *IoU* score obtained by ARTv2 foot detector highlights the potential benefits of using ARTv2 to automatically annotate and record gait information for videos. It also has the benefit of reducing the time needed to annotate a full video manually with software such as Supervisely (Supervise.ly, 2019) or MWA (Hewitt et al., 2016). The *IoU* score for the foot detector is relatively low, which could be due to a number of reasons. This includes the ground truth bounding boxes lying outside the area of the body, and the number of parameters used to fine-tune the foot tracking affecting the *IoU* score. Finally, *IoU* treats errors the same in small bounding boxes and large bounding boxes. A small error in a large predicted or ground truth bounding box is generally benign but a small error in a small predicted or ground truth bounding box has a much greater effect on the *IoU* score. This could

also explain why the *IoU* values for mouse were much lower for the rat data. However, the difference in *IoU* between rat and mouse could also be due to the settings creating artefacts, or erroneously eroding features needed to identify feet.

Locowhisk System outputs

The mean statistics of the outputs from ARTv2, are consistent with data previously recorded for rat and mice. The mean whisker frequency for rats (11.04 ± 5.08 Hz) and mice (19.49 ± 8.02 Hz) are similar to what was recorded in Mitchinson et al. (2011) (8.55 ± 0.62 Hz and 11.35 ± 0.95 Hz, respectively) and Grant et al. (2018a) (8.8 ± 0.76 Hz and 16.08 ± 7.05 Hz, respectively). The protraction speed recorded in ARTv2 for both species (1.434 ± 0.51 deg/ms and 1.535 ± 0.79 deg/ms) is very similar to values recorded in Grant et al. (2018a) (1.34 ± 0.1 and 1.69 ± 0.31, respectively). However the retraction speeds of rats and mice (1.274 ± 0.51 deg/ms, 1.535 ± 0.79 deg/ms respectively) are higher than those recorded in Grant et al. (2018a) (0.15 ± 0.05 deg/ms, 0.43 ± 0.23 deg/ms, respectively). The whisker amplitude for rats and mice (24.2± 9.30˚, 17.36 ± 4.47 ˚ respectively) is much smaller than values in Grant et al. (2018a), yet is similar to the model calculated amplitude (29.19 ± 7.14 ˚, 16.21 ± 6.49 ˚, respectively) in Mitchinson et al. (2011). The mean whisker angles recorded in ARTv2 were very similar in both species. However, Mitchinson et al. (2011) and Grant et al. (2018a), found that the mouse whisker angle (or offset) is often lower in rats. Differences in whisker measurements, could be due to individual differences in both whisker movements or locomotion, as whisker movements and positioning is associated with locomotion speed (Arkley et al. 2014).

Results from the ARTv2 foot detector are also similar to data from the literature. The recorded mean stride distance for rats and mice for fore (101.5 ± 42.0 mm, 56.95 ± 10.3 mm) and hind (48.52 ± 0.0 mm, 51.42 ± 13.93 mm) feet respectively in ARTv2, are very similar to fore feet data from Santos et al. (1995) of 113 mm for rats and mice of 66-70 mm (Amende et al., 2005; Virmouni et al., 2014). The animals in this study were filmed in a free and open environment, rather than a treadmill or trackway; therefore, there is probably some differences in gait, especially in the timing. Indeed, rats in this study have slightly longer stance (fore: 302.0 ± 89.75 ms, hind: 365.4 ± 248.7 ms) and swing times (385.2 ± 130.8 ms) for their forefeet, compared to data from Santos et al. (1995) (stance: 200ms, swing:130 ms) and Herold et al. (2016). The mouse stance times (fore: 131.7 ± 34.38 ms, hind: 171.3 ± 72.28 ms) are similar to those recorded in Mendes et al., (2015). Although the swing phase of the mice feet in this study are slightly longer (fore: 498.3 ± 374.0 ms, hind: 216.3 ± 40.42 ms) than those recorded in Mendes et al., (2015). This could be due to a number of factors, such as the animals exploring and interacting with elements within the LocoWhisk arena, such as objects and walls. Measurements of gait recorded in the literature are usually conducted either on a treadmill, trackway or in tunnel. These approaches do not necessarily represent coordinated movements and locomotion seen in freely moving and exploring rodents (Herbin et al., 2007; Batka et al., 2014; Guillot et al., 2008). Therefore, the LocoWhisk arena and software (ARTv2) might encourage more natural locomotion behaviours and gaits than some other experimental set-ups.

The whisker and foot tracking algorithms incorporated into ARTv2, have been designed to be automated and reduce manual intervention while maintaining accuracy levels similar to that of other methods, particularly on datasets with low spatial resolution. ARTv2, has been designed to be usable by researchers who have little to no programming expertise, and is freely available as an executable file with a Graphical User Interface (GUI) at: https://github.com/AIEMMU/ARTv2. The source code is also available for researchers to reproduce the results in this paper, explore the code base and expand upon our current work (from: https://github.com/AIEMMU/ARTv2).

Conclusion

We demonstrate within this paper that the LocoWhisk arena and ARTv2 are able to detect feet and whiskers accurately. The LocoWhisk system and ARTv2 represent the first step in developing a robust, fully-automated whisker and foot detector for detecting whiskers and feet in freely moving rodents in a range of experimental conditions, including open-field and during object exploration. The whisker detection algorithm has been developed as an alternative to existing whisker detectors. It has been designed to reduce as much manual intervention as possible, while maintaining accuracy levels similar to that of other methods. This was achieved by dynamically calculating each required setting per frame, and can thus be used on datasets consisting of various brightness and contrast changes. The foot tracking algorithm has been developed with a similar ethos to reduce user interaction, and to be usable on other datasets of varying brightness. Both achieve robust and accurate detection compared to manual tracking. Future software development may include incorporating machine learning algorithms, or integrating behavioural tracking with electrophysiology or electromyography measurements. This would enable musculoskeletal or neural measurements to be collected alongside behavioural assessments.

The LocoWhisk arena can robustly measure rodent whisker movements and gait. It can be applied to rodent models to help us to better understand how disease progression can affect motor control, which may lead to the development of new rodent models and the design of new treatments.

Acknowledgements

We would like to thank Antonello Mallamachi and Marco Stebel for providing access to the opossum colony at the University of Trieste animal facility, and to Hazel Ryan and Vicki Breakell for their continued support and access to animals at the Wildwood Trust. We are grateful to Tony Prescott, Ben Mitchinson and Kendra Arkley at the University of Sheffield for their support producing the pilot set-up, which first combined whisker filming with a pedobarograph.

References

Ahissar, E. and Knutsen, P. M. (2008) 'Object localization with whiskers.' *Biological cybernetics*, 98(6) pp. 449-458.

Arkley, K., Grant, R. A., Mitchinson, B. and Prescott, T. J. (2014) 'Strategy change in vibrissal active sensing during rat locomotion.' *Current Biology*, 24(13) pp. 1507-1512.

Bale, M. R., Campagner, D., Erskine, A. and Petersen, R. S. (2015) 'Microsecond-scale timing precision in rodent trigeminal primary afferents.' *Journal of Neuroscience*, 35(15) pp. 5935-5940.

Bradski, G. and Kaehler, A. (2008) *Learning OpenCV: Computer vision with the OpenCV library.* " O'Reilly Media, Inc.".

Clack, N. G., O'Connor, D. H., Huber, D., Petreanu, L., Hires, A., Peron, S., Svoboda, K. and Myers, E. W. (2012) 'Automated tracking of whiskers in videos of head fixed rodents.' *PLoS computational biology*, 8(7) p. e1002591.

Corr, S., McCorquodale, C. and Gentle, M. (1998) 'Gait analysis of poultry.' *Research in veterinary science*, 65(3) pp. 233-238.

Dauchy, R. T., Wren, M. A., Dauchy, E. M., Hoffman, A. E., Hanifin, J. P., Warfield, B., Jablonski, M. R., Brainard, G. C., Hill, S. M. and Mao, L. (2015) 'The influence of red light exposure at night on circadian metabolism and physiology in Sprague–Dawley rats.' *Journal of the American Association for Laboratory Animal Science*, 54(1) pp. 40-50.

Devernay, F. (1995) *A non-maxima suppression method for edge detection with sub-pixel accuracy.* INRIA.

Garland, H., Wood, N. I., Skillings, E. A., Detloff, P. J., Morton, A. J. and Grant, R. A. (2018) 'Characterisation of progressive motor deficits in whisker movements in R6/2, Q175 and Hdh knock-in mouse models of Huntington's disease.' *Journal of neuroscience methods*, 300 pp. 103-111.

Gonzalez, R. C. and Woods, R. E. (2002) 'Thresholding.' *Digital Image Processing*, pp. 595-611.

Grant, R. A., Mitchinson, B. and Prescott, T. J. (2012) 'The development of whisker control in rats in relation to locomotion.' *Developmental psychobiology*, 54(2) pp. 151-168.

Grant, R. A., Breakell, V. and Prescott, T. J. (2018a) 'Whisker touch sensing guides locomotion in small, quadrupedal mammals.' *Proceedings of the Royal Society B: Biological Sciences*, 285(1880) p. 20180592.

Grant, R. A., Mitchinson, B., Fox, C. W. and Prescott, T. J. (2009) 'Active touch sensing in the rat: anticipatory and regulatory control of whisker movements during surface exploration.' *Journal of neurophysiology*, 101(2) pp. 862-874.

Grant, R. A., Haidarliu, S., Kennerley, N. J. and Prescott, T. J. (2013) 'The evolution of active vibrissal sensing in mammals: evidence from vibrissal musculature and function in the marsupial opossum Monodelphis domestica.' *Journal of Experimental Biology*, 216(18) pp. 3483-3494.

Grant, R. A., Wong, A. A., Fertan, E. and Brown, R. E. (2018b) 'Whisker exploration behaviours in the 5xFAD mouse are affected by sex and retinal degeneration.' *Genes, Brain and Behavior*, p. e12532.

Grant, R. A., Sharp, P. S., Kennerley, A. J., Berwick, J., Grierson, A., Ramesh, T. and Prescott, T. J. (2014) 'Abnormalities in whisking behaviour are associated with lesions in brain stem nuclei in a mouse model of amyotrophic lateral sclerosis.' *Behavioural brain research*, 259 pp. 274-283.

Grant, R. A., Cielen, N., Maes, K., Heulens, N., Galli, G. L., Janssens, W., Gayan-Ramirez, G. and Degens, H. (2016) 'The effects of smoking on whisker movements: A quantitative measure of exploratory behaviour in rodents.' *Behavioural processes*, 128 pp. 17-23.

He, K., Gkioxari, G., Dollár, P. and Girshick, R. (2017) *Mask r-cnn.*

Herold, S., Kumar, P., Jung, K., Graf, I., Menkhoff, H., Schulz, X., Bähr, M. and Hein, K. (2016) 'CatWalk gait analysis in a rat model of multiple sclerosis.' *BMC neuroscience*, 17(1) p. 78.

Hewitt, B., Yap, M. H. and Grant, R. A. (2016) 'Manual whisker annotator (mwa): A modular open-source tool.' *Journal of Open Research Software*, 4(1)

Hewitt, B. M., Yap, M. H., Hodson-Tole, E. F., Kennerley, A. J., Sharp, P. S. and Grant, R. A. (2018) 'A novel automated rodent tracker (ART), demonstrated in a mouse model of amyotrophic lateral sclerosis.' *Journal of neuroscience methods*, 300 pp. 147-156.

Kloefkorn, H. E., Pettengill, T. R., Turner, S. M., Streeter, K. A., Gonzalez-Rothi, E. J., Fuller, D. D. and Allen, K. D. (2017) 'Automated Gait Analysis Through Hues and Areas (AGATHA): a method to characterize the spatiotemporal pattern of rat gait.' *Annals of biomedical engineering*, 45(3) pp. 711-725.

Knutsen, P. M., Derdikman, D. and Ahissar, E. (2005) 'Tracking whisker and head movements in unrestrained behaving rodents.' *Journal of neurophysiology*, 93(4) pp. 2294-2301.

Lin, T.-Y., Goyal, P., Girshick, R., He, K. and Dollár, P. (2017) *Focal loss for dense object detection.*

Mendes, C. S., Bartos, I., Márka, Z., Akay, T., Márka, S. and Mann, R. S. (2015) 'Quantification of gait parameters in freely walking rodents.' *BMC biology*, 13(1) p. 50.

Mitchinson, B., Grant, R. A., Arkley, K., Rankov, V., Perkon, I. and Prescott, T. J. (2011) 'Active vibrissal sensing in rodents and marsupials.' *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1581) pp. 3037-3048.

O'Connor, D. H., Clack, N. G., Huber, D., Komiyama, T., Myers, E. W. and Svoboda, K. (2010) 'Vibrissa-based object localization in head-fixed mice.' *Journal of Neuroscience*, 30(5) pp. 1947-1967.

Pammer, L., O'Connor, D. H., Hires, S. A., Clack, N. G., Huber, D., Myers, E. W. and Svoboda, K. (2013) 'The mechanical variables underlying object localization along the axis of the whisker.' *Journal of Neuroscience*, 33(16) pp. 6726-6741.

Peirson, S. N., Brown, L. A., Pothecary, C. A., Benson, L. A. and Fisk, A. S. (2018) 'Light and the laboratory mouse.' *Journal of neuroscience methods*, 300 pp. 26-36.

Perkon, I., Košir, A., Itskov, P. M., Tasič, J. and Diamond, M. E. (2011) 'Unsupervised quantification of whisking and head movement in freely moving rodents.' *Journal of Neurophysiology*, 105(4) pp. 1950-1962.

Real, R. and Vargas, J. M. (1996) 'The probabilistic basis of Jaccard's index of similarity.' *Systematic biology*, 45(3) pp. 380-385.

Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. (2016) *You only look once: Unified, real-time object detection.*

Rosenfeld, A. (1976) *Digital picture processing.* Academic press.

Sofroniew, N. J., Cohen, J. D., Lee, A. K. and Svoboda, K. (2014) 'Natural whisker-guided behavior by head-fixed mice in tactile virtual reality.' *Journal of Neuroscience*, 34(29) pp. 9537-9550.

Supervise.ly. (2019) *The leading platform for entire computer vision lifecycle: Iterate from image annotation to accurate neural networks 10x faster*. [Online] [Accessed on 01/02/2019]

Technology, N. I. (2012) CatWalk XT Version 10.0. Reference Manual. Noldus Information Technology b.v. , Wageningen, The Netherlands.

Voigts, J., Sakmann, B. and Celikel, T. (2008) 'Unsupervised whisker tracking in unrestrained behaving animals.' *Journal of neurophysiology*, 100(1) pp. 504-515.

FIGURE CAPTIONS

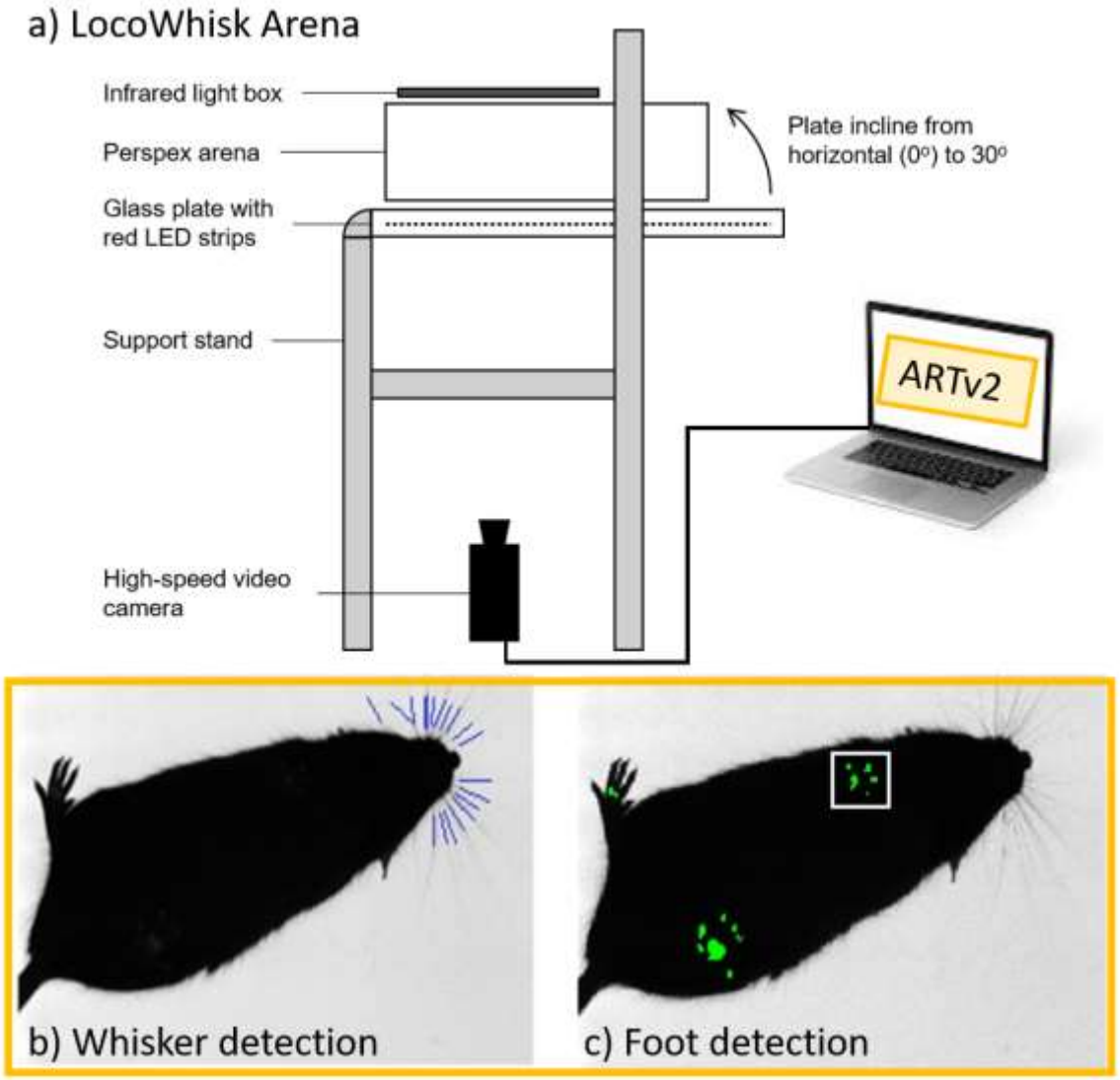

Fig. 1: The LocoWhisk Arena and ARTv2 software. (a) schematic of arena; (b) Inset: Example high-speed video footage, showing overlaid whisker detection; and (c) foot detection.

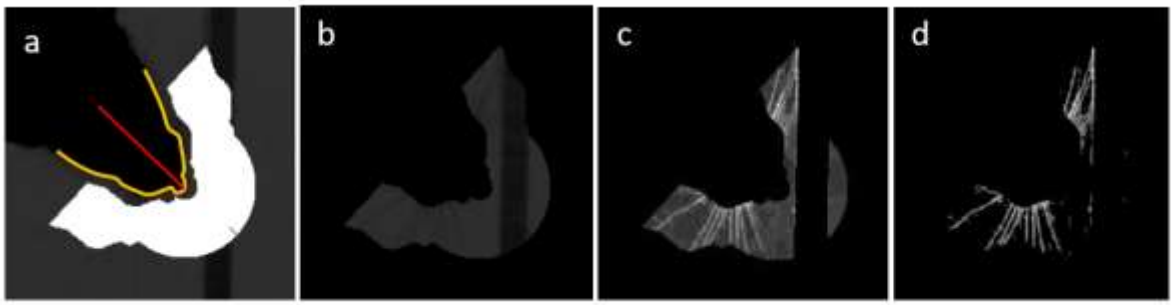

Fig. 2: Illustration of the ROI image enhancement stages. a) The computed ROI mask; b) ROI extraction; c) Linear Normalisation; and d) Threshold to zero.

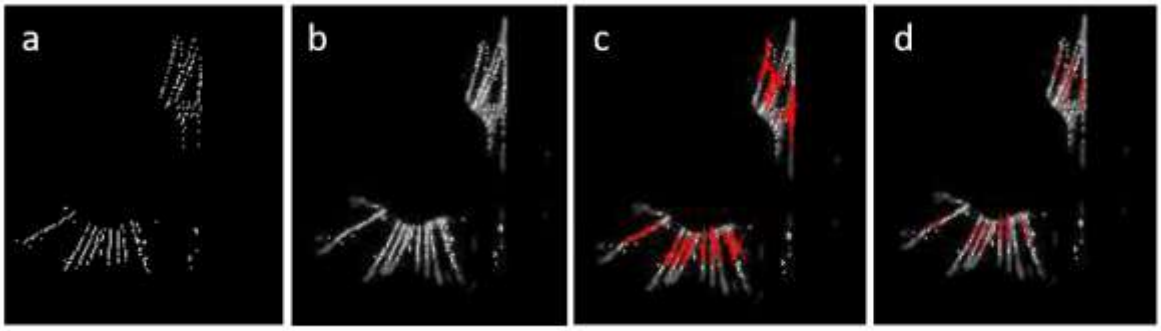

Fig. 3: An overview of the whiskers localisation and detection. a) Probable whisker locations; b) Probably whisker locations added to the ROI image; c) All potential whisker lines; and d) The detected whiskers.

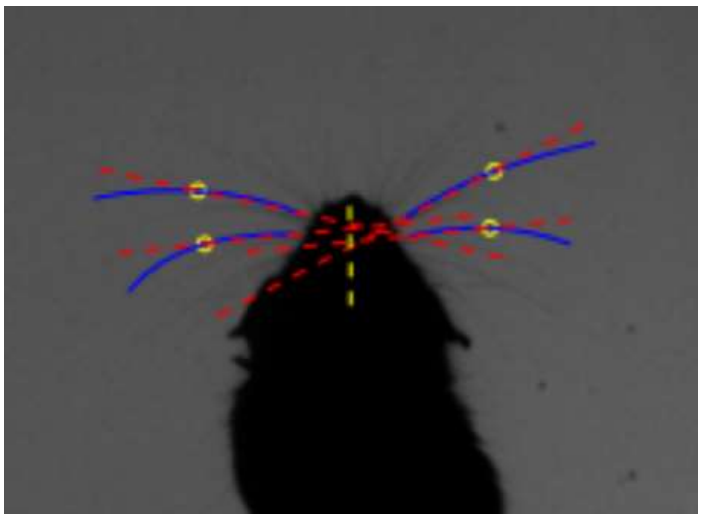

Fig. 4: Whisker angular positions. The whisker angle is calculated by taking the tangent from the mid line of the rodent's head to a point on a given whisker shaft.

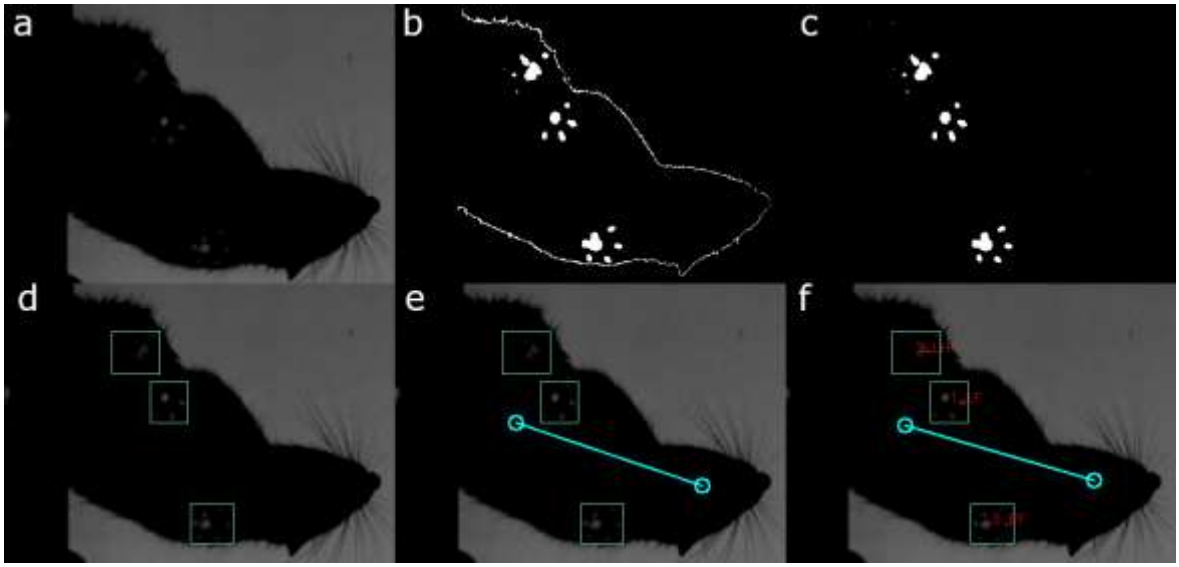

Fig. 5: Visual illustration of the proposed foot detection algorithm. a) Input frame from a video; b) Mask created by segmenting the body from the image and converting to the log colour space; c) Eroded mask; d) Foot bounding boxes; e) Bounding boxes and the line splitting the rodent; f) Correctly labelled feet.

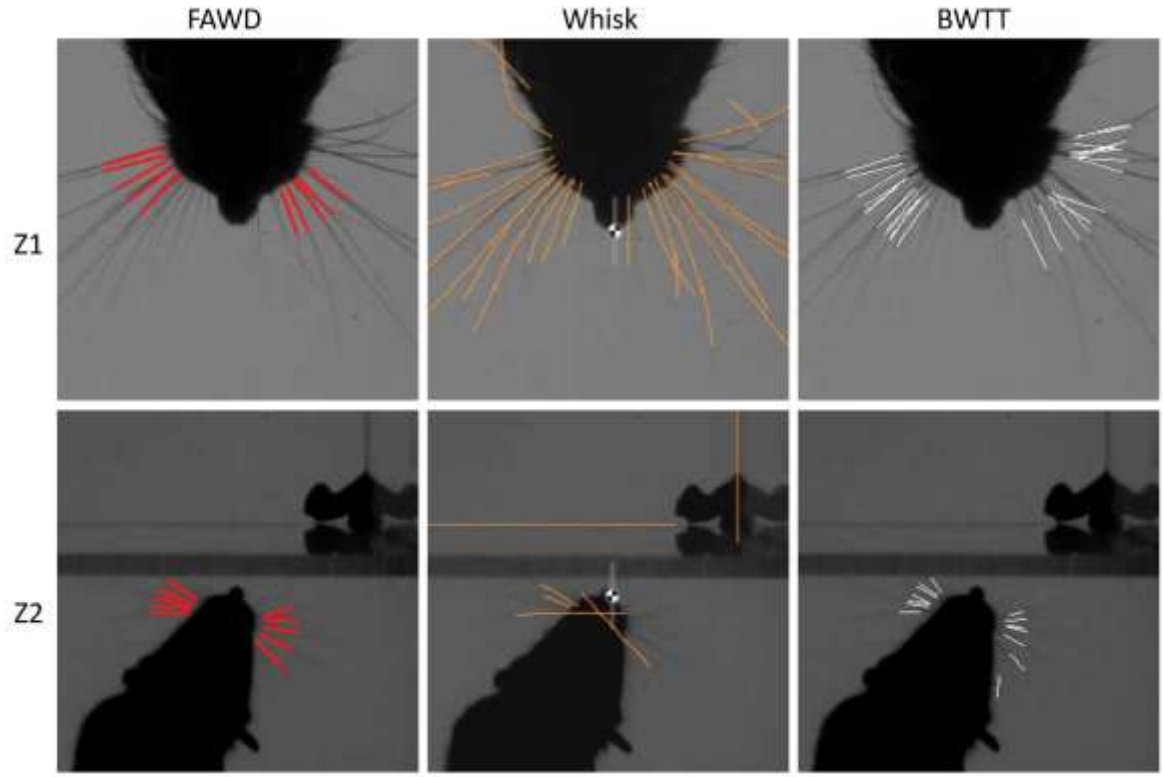

Fig. 6: Comparison of different whisker detectors on two different Zoom levels. The first column is Artv2, the 2nd is Whisk and the 3rd is BWTT. The first row (a) is a higher-resolution dataset; the Bottom: b) is a lower resolution dataset.

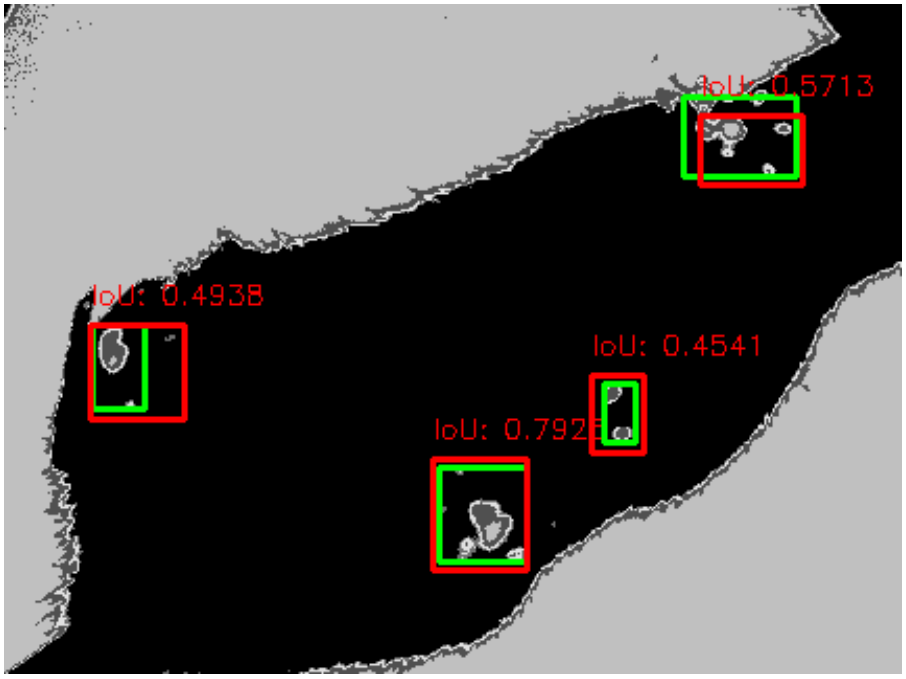

Fig. 7: An example of image, with the predicted *IoU* score for each foot print. Green boxes are ground truth annotations and red boxes are predicted boxes.

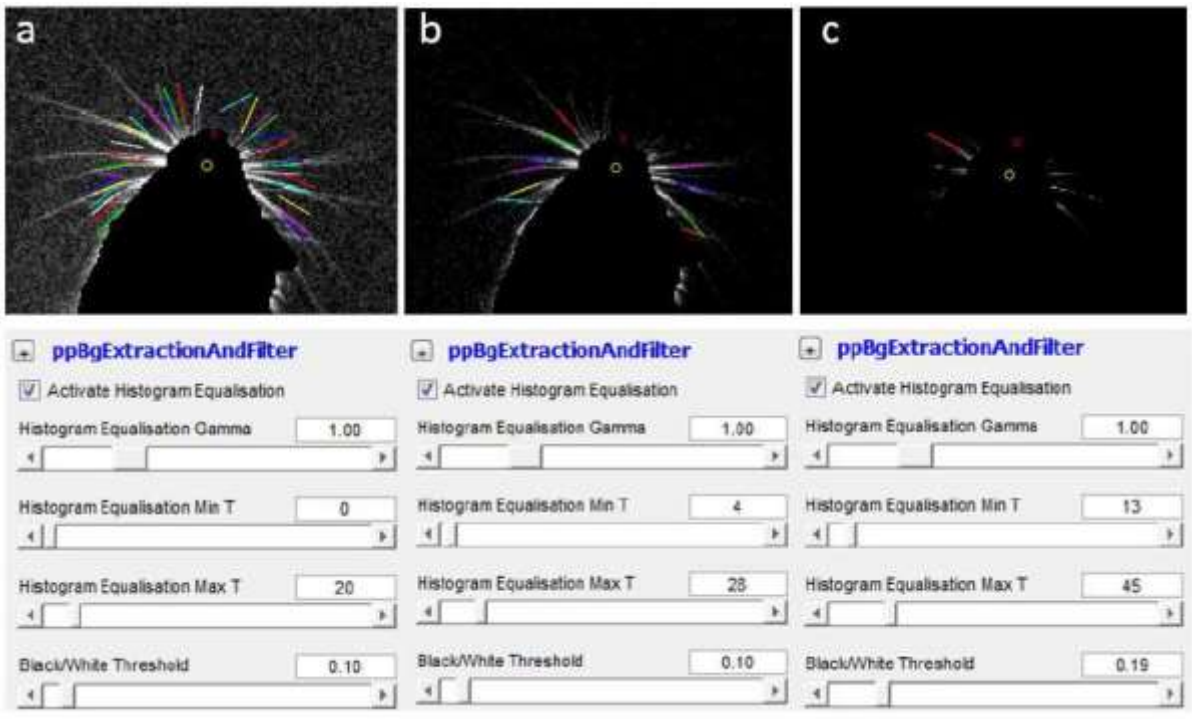

Fig. 8: Variation in detection rates for different settings in BWTT