

# MARGO Paper - Draft

September 25, 2018

# 1 Abstract

## 2 Introduction

Improvements in the accessibility and efficiency of algorithms, computer hardware, and electronic sensors have brought the age of omics to bear on the study of animal behavior. The statistical power afforded by the large sample sizes possible with automated methods make it invaluable in dissecting out the various determinants of behavior as manual annotation of video data is cumbersome and subject to individual bias. Automated animal tracking methods have become commonplace in behavioral study and vary in computational complexity and specialization to particular imaging configurations. Some existing platforms are designed to track and maintain identities of individuals in mixed groups individuals [refraf]. Improvements in machine learning approaches to object localization and classification make it possible to train models not only to accurately track a variety of animal species but even to distinguish and maintain identities in mixed groups [refraf]. Tracking identity in mixed groups requires resolving identities through collisions where bodies are overlapping. This approach offers the ability to study complex social and individual behaviors, but the computational cost of collision resolution means that tracking must be performed offline on recorded video data [refraf]. These methods can be rate-limiting for massively high-throughput experiments testing large numbers of experimental groups. Furthermore, the need to record uncompressed or lightly compressed high-resolution video data and the propagation of errors (even at low rates) make it impractical for continuous tracking over timescales of days or weeks.

In contrast, real-time tracking offers the benefits of closed-loop stimulus delivery and a small data footprint. Previously described platforms for real-time tracking offer a variety of features including tracking of single or mixed groups of individuals, modular arena design, and closed-loop stimulus delivery [refraf]. Several such platforms or software packages have been published to perform a variety of real-time tracking tasks [refraf]. The tracking algorithms, software interface, hardware configurations, and experimental goals of available trackers vary greatly. Existing packages span the range of stripped-down APIs for implementation of a single algorithm to graphical user-interfaces that capable of defining custom experiments and tracking [refraf]. We wanted to develop a platform that could integrate many of these features into a single software package and further build on them to provide a powerful interface for defining custom high-throughput experiments with minimal assumptions about the configuration of the experiment. Additionally, we wanted to build a package with an emphasis on simple, fast tracking for very large numbers

of individuals or experimental groups over long time-scales.

We developed MARGO, a MATLAB based tracking suite, with these concerns in mind. We found that this configuration can reliably track large cohorts of individuals or distinct experimental groups in real-time over very long timescales. Due to its flexible tracking algorithms, it can be used to track a wide variety of organisms as groups of isolated individuals or mixed populations over varying imaging conditions. Additionally, we demonstrate that integration of peripheral hardware into our tracking suite enables closed-loop individual stimulus delivery in high-throughput paradigms. We packaged MARGO with an easy-to-use graphical user interface (GUI) and comprehensive user manual to minimize the setup time required to use it. We also demonstrate that tracking requires little to no special hardware and works reliably with minimalist configurations of commonly available items. Taken together with its small number of configurable parameters and good visual feedback on tracking configurations, we believe that MARGO is ideal for large behavioral screens, experiments requiring real-time stimulus delivery, and users looking to run rapid pilot experiments with little setup.

## Results

### *MARGO Description*

In designing MARGO, we decided that it should 1.) track real time at a high acquisition rate to facilitate tight closed-loop, individual stimulus delivery 2.) use a strategy capable of tracking maintaining identities of hundreds to thousands of individuals with a single camera or video 3.) work for many organisms, behavioral platforms, and imaging configurations 4.) be robust to imaging noise and perturbation to facilitate long-term imaging over days or weeks and 5.) have a minimal data footprint to enable large behavioral screens.

To address the first two design constraints, we used a simplifying assumption that each tracked object or experimental group occupies a distinct region of interest (ROI). This assumption offers the advantage of maintaining individual identities indefinitely without the need to manually resolve collisions and facilitates accurate tracking with minimal assumptions about the size and morphology of tracked objects. This fact dramatically improves the throughput of large experiments where accurately maintaining identity is important not only by limiting tracking time to the duration of acquisition but also by drastically reducing the requirements for human supervision and

intervention. We found that insisting on ROI definition ultimately relaxes the computational requirements enough that thousands of individuals can plausibly be tracked in real-time. In cases where maintaining individual identity is unimportant, such as population assays, the ROI-based architecture of MARGO makes it simple to test multiple populations simultaneously in separate areas of the same field of view. By assigning each tracked object to an ROI MARGO can track hundreds of individuals in each arena without swapping population identities.

For these reasons, MARGO setup is organized around a core four-step tracking workflow: 1.) define ROIs, 2.) collect a background reference image, 3.) sample imaging statistics for noise detection and correction, 4.) perform tracking [fig.1]. To fully take advantage of the throughput offered by tracking thousands of ROIs, we insisted on strategies of ROI definition that would make it simple to rapidly define thousands of ROIs while maintaining the flexibility necessary to accommodate a wide variety of experimental configurations. To address these concerns, MARGO initiates tracking setup by defining ROIs through one of two modes. The first is an automated detection mode that detects and segments regular patterns of high-contrast regions in the image. The second mode where the user manually can place grids of ROIs of any arbitrary dimensions with each ROI having a quadrilateral or ellipsoid shape. Following ROI definition, tracking is performed on difference images computed by subtracting the current frame from the background. Difference images are then segmented into foreground and background regions. Candidate blobs in the foreground are filtered by size and sorted to ROIs by spatial location. ROIs with excess candidates are assigned the best candidates based on size and position.

Degradation of difference image quality over time due changes in the background, noisy imaging, and physical perturbation of the imaging setup constitutes a significant barrier to long term tracking [refraf]. The sensitivity of difference images to minor frame to frame changes makes them particularly well suited to low-resolution tracking on complex backgrounds but also exposes the tracking to noise introduce by small changes in the image background such as misalignment or changes in illumination. To address this problem, MARGO continuously monitors the quality of the difference image updates or reacquires the background reference when imaging becomes noisy. Prior to tracking, the software samples the distribution of the total number of above-threshold pixels under clean imaging conditions to serve as a baseline for comparison. During tracking, the software then continuously calculates statistics on a similar distribution acquired on a rolling basis. Re-referencing is triggered when the rolling sample significantly deviates from the baseline distribution.

*MARGO is robust to noise*

We assessed the ability to MARGO to handle degradation of the difference image by repeatedly shifting the background reference by a small amount to simulate misalignment caused by physical perturbation. Tracking was performed live on unaltered image data to serve as a ground truth while simultaneously recording video data. Background misalignment was repeatedly triggered on a fixed interval by shifting the background reference by 2 pixels in a random direction. Trial triggered averaging of the tracking error shows that MARGO reliably detects the misalignment and fully recovers clean tracking quickly [fig.2]. Forcing reacquisition of the background reference has the disadvantage of initializing the reference with a single image, meaning that a median reference image cannot be computed until subsequent references are acquired. We found that this typically causes a reduction in tracking accuracy that is brief (<2s) and had little effect on the overall integrity of the full trace. Indeed we found a negligible effect on tracking error (3.07+/- 2.5) even when shifting the background every 2 seconds.

We further tested the robustness of MARGO to noisy imaging by artificially injecting noise into the threshold image of each frame of a video previously acquired and tracked under clean conditions. We injected noise into each frame by randomly setting each pixel to True with a uniform probability. Noise was added to the binary image downstream of noise correction and upstream of tracking to simulate tracking under conditions where noise detection and correction are poorly calibrated. Tracking error was measured as the deviation in tracked positions between each frame of the video with added noise and the unaltered source video. Even in the absence of noise correction, we measured sub-pixel level error up to 20% noise. In practice, we find it easy to create imaging conditions with noise levels well below the levels tested without the use of sophisticated tracking setups.

*MARGO is ideal for massive screens*

We designed MARGO with high-throughput behavioral screens in mind, where hundreds or thousands of animals may need to be tested from hundreds of experimental groups each. For this reason, many features of MARGO have been included to reduce the time needed to establish successful tracking. We found that the ability to receive real-time visual feedback on parameters, features for saving and loading GUI configurations, the absence of model training, and automated ROI detection all substantially reduce the time needed to setup tracking. The speed of the tracking algorithm also makes it possible to track very large numbers of animals simultaneously in a single field of view, meaning that multiple experimental groups can be tested simultaneously by a single machine. The mean acquisition rate was measured while systematically varying the number of tracked ROIs to assess MARGO's capacity for massively high-throughput experiments. We found that

the time complexity of the core tracking routine scaled linearly as a function of the number of ROIs tracked [fig. 2]. On modern computer hardware (computer specs), we measured acquisition rates above 5Hz up to 2400 ROIs and reached as high 160Hz for single ROI tracking. Achieving an appropriate frame rate will of course depend on the timescale of the targeted movement dynamics. We found that frame rates above 5Hz were more than sufficient to capture features of single animal movement bouts in adult fruit flies.

### *MARGO is simple and flexible*

Both the ease of setup and the flexibility of the software make MARGO an ideal platform for running experiments. To demonstrate the simplicity of prototyping experiments in MARGO without the need for specialized hardware or imaging configurations, we ran a tracking experiment under minimal conditions using only commonly available materials. Individual fruit flies were placed in a 48 well culture plate. The plate was put in a cardboard box with a sheet of white paper on the floor of the box to reduce reflections and improve contrast. Movies were recorded on a 1.3MP smartphone camera and imported into MARGO. Tracks and movement bout features acquired under minimal conditions showed no apparent differences to those acquired under optimized conditions. We found that minimal conditions narrowed the functional range of parameters adjusted during setup and slightly lowered the acquisition rate by introducing noise but had no apparent effect on the accuracy of traces.

MARGO was developed for high-throughput ethology in fruit flies. Many small organisms used for high-throughput behavior, such as nematodes, larval zebrafish, and fruit fly larvae are more translucent than adult flies. To assess MARGO's robustness to organisms and lower contrast imaging, we tested tracking performance on sample videos from a variety of common model organisms. Higher translucence expectedly narrowed the functional range of some parameters, but we found that real-time feedback on parameter performance made it simple to adjust parameter values to the new organisms and imaging conditions. Sample traces acquired from other organisms were qualitatively similar to those acquired with adult flies, suggesting that MARGO's high-throughput, ROI-based tracking approach is flexibly applicable to a variety of organisms.

We integrated the functionality of MARGO into a graphical user interface (GUI) to make the full functionality of the software accessible to users unfamiliar with MATLAB or programming in general [fig. 1]. We generally find that new users easily learn to use both the work-flow and parameter customization. Furthermore, live feedback on the effects of parameter and option setting make it easy to configure tracking for a variety of experimental paradigms. We found that typical setup of tracking ranged between

a few seconds to a few minutes. The utility of the GUI extends beyond tracking setup to include customization of analysis, visualization, and input/output sources such as videos, cameras, displays, and COM devices. A user's manual is additionally included with complete details on using the GUI, setting up tracking, utilizing the data output, and defining custom experiments via the application programming interface (API).

Behavioral experiments are frequently more complex than tracking objects in a dish. Custom experiments routinely require complex arena geometries, data streams from external sensors, control of peripheral hardware, and time-series data of various object features. MARGO offers a set of features that make it particularly adept at defining new behavioral paradigms[tab]. MARGO was designed to run experimental modules that define how to coordinate the various activities of custom experiments. The ROI definition strategies employed here are robust to many arena geometries with minimal constraint on the positioning or regularity of ROIs within the field of view. MARGO provides support for external hardware detection and configuration. Data output is also easily configurable and customizable. Templates for new tracking routines can be automatically generated and detected from within the GUI. In practice, we find that new experiments with customized output and hardware interfaces can typically be defined in 1-2 simple functions with some knowledge of the API. Taken together, these features make MARGO well-suited to new behavioral modules with distinct organisms, arena configurations, stimulus epochs, and data outputs.

### *Closed-loop hardware control with MARGO*

Real-time tracking opens up the possibility of closed-loop stimulus delivery conditional upon the behavior animals. MARGO offers native support of cameras for real-time image acquisition, displays for visual stimulus delivery, and serial COM devices digital input/output of peripheral electronics. The ubiquitous spread and adoption of easily-configurable consumer microcontrollers makes it simple to control a wide variety of devices. For this reason, MARGO was designed to detect and communicate with configurable COM devices to integrate real-time feedback from external sensors and coordinate closed-loop control of electronic hardware.

We modified an established behavioral platform for measuring locomotor bias of fruit flies (Y-maze assay) to include a phototactic choice in parallel [ref]. In the unmodified assay, a turn bias score is defined by recording the fraction of trajectories through the center of the maze followed by a turn to the right. To add a phototactic stimulus to the platform, a custom printed circuit board was used to implement independent control of white LEDs positioned at the end of each arm of the maze [fig. 4]. For a single trial, an LED was randomly lit in either the left or right unoccupied arms. Using a closed-loop between camera and the custom

circuit board, trajectories made from one arm of the maze triggered a presentation of a new phototactic choice. Once a change in arm was detected a new trial was initiated by randomly selecting an LED in the unoccupied arms to turn on. This process repeats over two hours until the experiment ends. A phototactic bias is scored by calculating the fraction of turns into the lit arm for each individual. Individuals typically make hundreds of choices over the course of an experiment, effectively decoupling the phototactic and locomotor choices [fig. 4]. Tiling many such mazes on a single board allowed the assay to be conducted at high-throughput. Overall, we recorded choices from (refraf) individuals, representing (refraf) choices in total.

Testing of a variety of wild-type strains in the LED Y-maze revealed a significant average positive photobias ( ) compared to that of blind flies (Norp-A) and flies tested with the LEDs off [fig. 4]. The lines tested showed considerable variation in mean population preference and individual variability. To assess the relationship between the phototactic and locomotor bias, all trials were subdivided by whether the light source appeared to the right or left of the starting point. We found that the mean turn bias but not the mean phototactic bias differed between these two conditions. Interestingly, categorizing trials this way revealed that the rank order of both phototactic bias and turn bias are anti-correlated between the two conditions, suggesting that phototactic and locomotor handedness affect the inform individual choices.

One benefit of high temporal resolution in real-time tracking is the ability to tightly and accurately modulate a stimulus in response changes in behavior. To test the ability of MARGO to deliver a stimulus dependent on accurate and rapid feedback on a frame-to-frame basis, we adapted a previously described optomotor paradigm to a high-throughput configuration in MARGO [refraf]. Under this paradigm, an optomotor response is elicited by targeting a high-contrast, rotating pinwheel directly beneath a freely walking fly. On average, this stimulus evokes a turn in the direction of the rotation to stabilize the visual motion [refraf]. Inaccurate targeting of the stimulus results in some disagreement in the direction of the visual motion. For this reason, it is important that the center of the pinwheel precisely and accurately follows the animal as it moves throughout the arena.

We constructed a behavioral platform with both a camera and an overhead mounted projector targeting an array of flat circular arenas. MARGO was used to create a mapping from the camera field of view to the projector's display. A closed-loop was established between the camera and the display was used to target stimuli independently to a population of individuals in parallel. To ensure faithful coordination between the tracking and stimulus delivery, the tracking rate was matched to the refresh rate of the display at 60Hz. Targeting stimuli this way, observed

that strong responses could only be reliably elicited when individuals were already moving prior to the presentation of the stimulus. We thus imposed an inclusive set of criteria for stimulus presentation. An individual was presented with a pinwheel stimulus when 1.) movement was observed 2.) a minimum inter-trial interval was exceeded, and 3.) a minimum distance to the edge of the arena was obtained. We included a minimum inter-trial interval to serve as a baseline for comparison to stimulus evoked behavior. A minimum edge distance was included to ensure that the stimulus occupied a significant portion of the animal's field of view.

Real-time applications Individual projector targeting Visual stimuli Optomotor Shade/light preference Individual optogenetic activation Real-time estimation of behavioral metrics LED Y-maze

### 3 Discussion

To the best of our knowledge, MARGO represents a uniquely powerful and flexible interface for real-time tracking, integrating many features of other previously described platforms such as: individual tracking, group tracking, arena modularity, small data footprint, and support for simple low-cost imaging configuration.

Furthermore, MARGO is uniquely specialized for crafting and delivering individually tailored stimuli under closed-loop control. The ability to rapidly define and track hundreds of thousands ROIs makes it simple to deliver stimuli separately. Taken together with built-in support for interfacing with hardware, MARGO is ideal for optogenetics and stimulus-driven ethology on a massive scale. Things to highlight in discussion?

Massively high-throughput support for experiments where groups are spatially segregated in separate ROIs. Groups can be single individuals or populations.

Simple, real-time, GUI-based tracking. No need for API to run a simple tracking experiment. Integrated support for cameras, displays, and serial hardware.

### 4 Methods

Methods Data and analysis Repository for data and analysis scripts Software The MARGO GUI, tracking algorithms, and all analysis software were written in MathWorks MATLAB. Detailed descriptions of the function and use of the MARGO GUI, ROI detection, tracking implementation, and noise correction are available in MARGO user manual. Optomotor stimuli were crafted and

displayed using the Psychtoolbox-3 for MATLAB. Software for control of all custom electronic hardware was written in C. Organism genotypes and care Flies - The genotype of all tested fruit flies was an inbred strain of Berlin K. Tracking experiments were performed with mixed groups of male and female flies 3-5 days post-eclosion unless otherwise noted. All animals were lightly anesthetized on ice and loaded into behavioral arenas. Animals were imaged and singly-housed on food in individual fly storage (FlySorter LLC) for all multi-day tracking experiments. Sample tracking experiments with other model organisms were all limited

to a single group each. C. Elegans were. Drosophila larvae. Larval zebrafish. All animals were tracked on food. Behavioral modules Unless otherwise specified, all tracking was conducted in custom imaging boxes constructed with acrylic siding and aluminum rails. Illumination was provided by dual-channel white and infrared LED array panels mounted at the base. Diffusers were placed in between the illuminator and the behavioral arena for smooth backlighting. All images were captured with a mounted overhead camera. Experimental procedures

#### Analysis and Statistics