

Система распознавания, отслеживания и описания поведения муравьев на видеозаписи, сделанной в полевых условиях

Е.Д. Московская¹✉, Е.В. Бургов¹, А.Д. Московский¹, Н.А. Гревцова¹

¹ Курчатовский комплекс НБИКС-природоподобных технологий,
НИЦ «Курчатовский институт»,
г. Москва, 123182, Россия

Ссылка для цитирования

Московская Е.Д., Бургов Е.В., Московский А.Д., Гревцова Н.А. Система распознавания, отслеживания и описания поведения муравьев на видеозаписи, сделанной в полевых условиях // Программные продукты и системы. 2024. Т. 37. № 1. С. 83–94. doi: 10.15827/0236-235X.142.083-094

Информация о статье

Группа специальностей ВАК: 2.3.1

Поступила в редакцию: 11.07.2023

После доработки: 21.08.2023

Принята к публикации: 05.09.2023

Аннотация. Статья посвящена разработке системы распознавания, отслеживания и оценки количественных характеристик поведения муравьев на видеопоследовательностях. Подобное ПО необходимо для биологических исследований, а конкретно – для изучения муравьев (мирмекология). Текущая версия системы ориентирована на расчет динамической плотности особей на территории – количества рабочих муравьев в заданной области за минуту. Рассматриваются видеозаписи, сделанные в полевых условиях на искусственной размеченной поверхности, для определения динамической плотности особей на территории. Для распознавания отдельных особей использован нейросетевой детектор, определяющий положение и ключевые точки муравья, соответствующие его голове и брюшку. По ключевым точкам рассчитывается ориентация объекта в пространстве. Точность определения границ изображений муравьев составила 92 %, а ошибка определения ориентации – 0.4 радиана. За счет определения ориентации объекта для решения задачи отслеживания используется расширенный фильтр Калмана, который учитывает ее в прогнозировании движения. Для решения задачи соотношения данных был применен критерий минимального времени перемещения муравья. Траектории движения муравьев, полученные в процессе работы системы, используются для расчета динамической плотности в произвольной области. Для отсеивания ложноположительных срабатываний детектора в системе отслеживания применяются разработанные эвристики. Благодаря используемым алгоритмам удалось добиться автоматического подсчета динамической плотности с точностью, не уступающей обработке видеозаписи человеком. Полученные результаты позволяют значительно облегчить работу мирмекологов с видеозаписями, заменив человеко-часы более дешевыми машино-часами. В дальнейшем система может быть дополнена блоками определения других количественных характеристик движения животных.

Ключевые слова: детектирование, отслеживание, муравьи, динамическая плотность особей, сверточная нейронная сеть, расширенный фильтр Калмана

Благодарности. Работа выполнена за счет государственного задания НИЦ «Курчатовский институт»

Введение. Статья посвящена разработке методов исследования поведения животных. В настоящее время наибольшую поддержку получают исследования, обеспеченные современной технической и методологической базами. В биологии, например, таковыми являются биохимические, генетические и др. Однако в изучении поведения мелких животных наблюдаются затруднения, поскольку эта область довольно слабо технически обеспечена.

Мирмекологические исследования (изучение муравьев) актуальны с точки зрения инженерной науки, так как муравьи являются модельным объектом для разработки биоподобных групп роботов в силу обладания рядом удобных параметров (социальность, относительная простота организации взаимодействия между индивидами в сравнении с человеком, разнообразие вариантов чувствительных органов и систем коммуникации и т.д.) [1]. Однако

эти работы весьма трудоемки из-за сложности макровидеосъемки и практически полного отсутствия доступных средств автоматического распознавания муравьев на видеозаписях. В данной статье подробно рассмотрены средства распознавания и их применение в мирмекологических исследованиях.

Как вспомогательный метод исследований весьма ценна и довольно широко применяется видеорегистрация. Она позволяет документировать наблюдения, собирать материалы, которые можно многократно обрабатывать, получая дополнительные данные [2, 3]. Однако использование видеорегистрации с обработкой видеозаписей вручную имеет ряд недостатков. При расчете некоторых параметров, например, динамической плотности особей на территории, время обработки видеозаписи вручную сопоставимо со временем получения аналогичных данных в поле (без видеорегистрации).

Для других параметров, например, для скорости движения муравьев, обработка видеозаписи вручную кропотлива и затратна по времени. В обоих случаях под вопросом необходимость применения видеорегистрации как таковой и целесообразность проведения исследований. Отдельные параметры невозможно или почти невозможно рассчитать без применения автоматических систем распознавания.

Рассмотрим подробнее расчет стандартного параметра – динамической плотности особей на территории, то есть количества муравьев на квадратном дециметре земли в минуту. Он необходим для оценки активности семьи муравьев на некоторой территории [4]. Показано, что от динамической плотности особей зависят многие особенности их поведения [5].

Для расчета в полевых условиях расчищается площадка, размечаются ее границы. Размеры площадки зависят от модельного вида муравьев. Наблюдатель располагается рядом с учетной площадкой, выжидает 5–10 минут (чтобы активность муравьев восстановилась после его появления) и в течение нескольких минут считает, сколько муравьев заходит на площадку. Потом рассчитывается значение динамической плотности особей на один квадратный дециметр в минуту.

При использовании видеорегистрации над учетной площадкой размещается камера. Далее видеозапись просматривается, ведется визуальный подсчет, но уже в лабораторных условиях. Находящееся в открытом доступе ПО не позволяет рассчитать динамическую плотность автоматически. Преимуществами видеосъемки являются документирование, более высокая точность расчета, возможность повторной обработки, меньшее воздействие на объект исследования. При работе в поле исследователь также может установить устройства видеорегистрации, а потом решать другие задачи. Однако перечисленные преимущества могут нивелироваться трудоемкостью работы со съемочной аппаратурой и обработкой видеозаписи. Кроме того, как уже говорилось, время обработки видеозаписи вручную сопоставимо со временем расчета динамической плотности особей в поле без использования каких-либо технических средств. В результате большинство отечественных мирмекологов рассчитывают динамическую плотность муравьев на территории в поле, даже не используя видеосъемку. Бывают также случаи отказа от расчета динамической плотности или других количественных параметров поведения муравьев из-за высокой тру-

доемкости. Для сокращения трудозатрат мирмекологов, а также для обеспечения этологических исследований необходимы разработка и внедрение методов автоматического распознавания муравьев на видеозаписях.

Обзор существующих подходов

Проблеме распознавания и отслеживания муравьев посвящено немало исследований.

Так, в работе [6] рассматривается система видеонаблюдения в реальном времени BACH, основанная на компьютерном зрении. По данным исследователей, точность слежения у BACH почти такая же, как у человека. В обнаружении объектов системой два этапа: первый – поиск объектов с помощью глубокой нейронной сети YOLO, второй – поиск маркеров ArUco, предварительно приклеенных к брюшку рабочих муравьев, анестезированных льдом. За обнаружением идет отслеживание, принцип работы которого состоит в сравнении вновь обнаруженных объектов со списком обнаруженных в предыдущем кадре за счет измерения перекрытий между их ограничивающими прямоугольниками. Оценка качества распознавания получилась неоднозначной: система хорошо распознает муравьев, которые двигаются медленно, однако имеет сбои тогда, когда муравей перемещается с большой скоростью или же попадет на нечеткий участок видео. Проблема слежения заключается в сложном движении муравьев: если объект внезапно меняет скорость, система не может соотнести его новое обнаружение с предыдущей траекторией и считает его новым объектом. Такая ситуация затрудняет последующий анализ поведения. Несмотря на недостатки, система используется в режиме реального времени, что может сэкономить время биологам. Однако самой трудоемкой задачей при использовании данного подхода является наклеивание меток на муравьев. Отсюда и главное ограничение подхода – невозможность проведения экспериментов в полевых условиях.

В исследовании [7] описана система обнаружения на основе глубокого обучения. Для поиска муравьев в каждом кадре использовалась глубокая сверточная нейронная сеть Mask R-CNN. Для получения хороших результатов распознавания модель обучена на большом объеме данных – 20 666 изображений. Обучение также было разделено на два этапа. На первом этапе модель делала первоначальные предсказания траекторий, а также давала подсказки

для выявления несоответствий в предсказанных траекториях муравьев и кадров, на которых муравьев трудно было обнаружить. На втором этапе к этим трудным случаям применялся отбор обучающих данных, чтобы найти дополнительные кадры для маркировки. Таким образом расширился набор обучающих данных для нового обучения модели. Хотя подход дал хорошие результаты, он потребовал очень больших затрат человеческих ресурсов на разметку. Отслеживание строилось на принципе оптимальной транспортировки (она же задача о назначениях, или коммивояжера), то есть оптимизации по перемещению всех муравьев между двумя последовательными кадрами. Такой метод, хотя и прост в реализации и в большинстве случаев верно составляет траектории, имеет сбои в моменты пересечения траекторий муравьев.

В школе информатики Сямэньского университета (Китай) был разработан метод обнаружения и отслеживания муравьев в помещении и на открытом воздухе [8–10]. Для обнаружения использовалась сверточная нейронная сеть ResNet-50, причем размер муравья считался фиксированным и равнялся 64×64 пикселя. Для отслеживания траекторий муравьев был внедрен фильтр Калмана. Однако авторы обращают внимание на то, что из-за сложного движения муравьев использование только фильтра Калмана не дает требуемой точности, поэтому добавляется дополнительный блок анализа – дескриптор внешнего вида. Это отдельная сверточная сеть, переводящая область изображения с муравьем в список признаков, которые впоследствии помогают идентификации объекта. Такой подход дал хорошие результаты при обработке видеозаписей, сделанных в лабораторных условиях. Во время съемки освещение не изменялось, положение камеры было жестко определено, что позволяло задать фиксированный размер изображения муравья на видеозаписи. На площадку, где велась съемка, не попадали насекомые других видов или другие объекты, которые могут быть ложно распознаны как муравьи. В полевых условиях работы мирмекологов почти невозможно или, по крайней мере, весьма трудозатратно соблюсти эти условия.

Постановка задачи

Система разработана для сокращения трудозатрат мирмекологов при расчете следующих количественных параметров поведения муравьев:

- динамическая плотность особей на территории;
- интенсивность движения на дорогах (количество муравьев, движущихся по дороге в одну и в другую стороны);
- индивидуальная скорость движения (средняя скорость по траектории, собственно скорость движения за вычетом времени остановок, иные варианты расчета скорости движения);
- автоматическое описание последовательностей действий с хронометражем (бег, остановка, бег, чистка, бег, контакт с другим муравьем и т.д.).

В качестве базовой задачи, на основании решения которой оцениваются перспективы применения данного подхода, используется расчет динамической плотности особей.

Для решения задач анализа поведения муравьев необходима информация о местоположении каждого муравья в конкретный момент времени, то есть требуется идентификация объектов для возможности отслеживания их между кадрами. Как видно из предыдущих исследований, использование одного фильтра Калмана не приводит к желаемым результатам из-за недостаточного количества данных и требует применения дополнительных механизмов. В настоящей работе предлагается расширить получаемые данные определением ориентации объектов, а также использовать расширенный фильтр Калмана, который способен ее учитывать.

Ориентация муравья в пространстве может быть рассчитана путем определения ключевых точек, соответствующих его голове и брюшку. Изображения головы и брюшка могут перекрывать друг друга, чаще всего это происходит при съемке сбоку. Чтобы минимизировать такие случаи, съемка должна вестись сверху.

В таблице 1 показаны способы детектирования муравьев в кадре, а также последующего трекинга в работах других авторов и в предлагаемом подходе.

Подход к разработке системы

Архитектура разработанной системы включает в себя три компонента: блок распознавания, блок отслеживания и блок анализа (рис. 1). Блок распознавания, построенный на нейросетевом подходе, представляет собой сверточную сеть поиска ключевых точек с предобученной магистралью. На вход системе распознавания подается кадр, а на выходе детектор выдает координаты ограничивающих рамок, ключе-

Таблица 1

Сравнение предлагаемого решения с существующими

Table 1

Comparing proposed solution with existing ones

Работа	Детектирование	Способ детектирования	Дополнительные ограничения
Предлагаемый подход	Keypoint-RCNN	Расширенный фильтр Калмана + минимально затраченное время перемещения	Съемка сверху
[6]	YOLO	На основе ArUco-меток и степени перекрытия ограничивающих рамок	Физическая разметка особей
[7]	Mask R-CNN	Оптимальная транспортировка	Большой набор данных
[8–10]	ResNet-50	Фильтр Калмана + визуальное сходство	Съемка в большом разрешении

вых точек, а также степени уверенности ограничивающих рамок. Для отслеживания множества объектов используется расширенный фильтр Калмана. На вход фильтру подаются данные, полученные от детектора, на выходе формируется список построенных траекторий. В данной работе система анализа реализует подсчет динамической плотности. Для определения этой характеристики достаточно иметь список траекторий с временными метками. На выходе система готовит график динамической плотности на выделенном участке.

Данные

Оригинальными данными являются видеозаписи муравьев, сделанные в полевых условиях. Частота кадров в данных видеозаписях находится в интервале 25–30 FPS. Разрешение исходных изображений – 1 920×1 080 пикселей. Исходный видеоматериал собирали в полевых условиях в Рязанской области в 2019–2020 гг. В качестве модельных видов использовали муравьев *Formica cinerea* и *Formica pratensis*, имеющих относительно крупных рабочих особей (около 0,5–1 см). Конечной целью сбора видеоматериала был расчет скорости движения рабочих муравьев. Методика позволяет также рассчитывать динамическую плотность особей. На территорию рядом с гнездом муравьев модельного вида устанавливали фанерную площадку размером 22×30 см, на которую наклеен лист акварельной бумаги с разметкой – квадратной сеткой с ячейками размером 5×5 см. Это позволяет определять реальные дистанции, которые преодолевают муравьи во время движения. Над площадкой располагали камеру так, чтобы в ее поле зрения попадала вся пло-

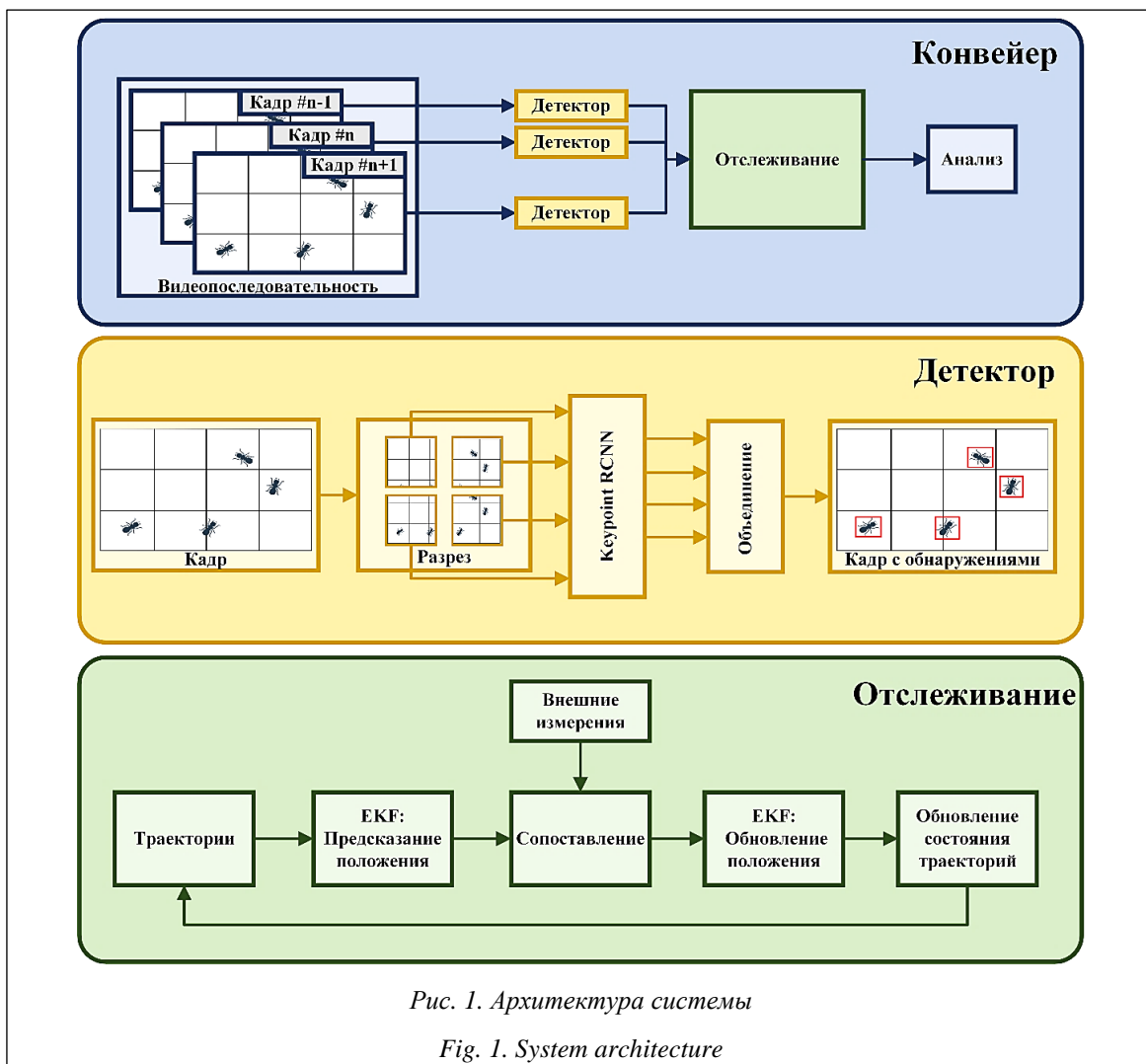
щадка. Далее на несколько часов запускали видеозапись. При необходимости (чтобы стимулировать движение муравьев) на территории или на самой площадке устанавливали сахарную кормушку (кусочек ваты, пропитанный сахарным сиропом).

Полученные записи обрабатывали при помощи видеоредактора. Из них выбирали фрагменты, позволяющие вести учет необходимых параметров поведения. Не использовали фрагменты, на которых к площадке подходил человек, осуществлялись манипуляции с камерой или с площадкой (и 5 минут после этого), а также присутствовали сильная вибрация штатива и, как следствие, камеры, резкие перепады освещения, при которых невозможно обнаружить муравьев на записи. На оставшихся видеофрагментах сохранялись небольшие вибрации камеры, а также перепады освещения, световые пятна, при которых можно распознать муравьев. Изображение кормушки, находящейся на площадке, закрывали цветным прямоугольником, чтобы не тратить ресурсы системы на распознавание и отслеживание муравьев, которые пьют сахарный сироп, поскольку это не нужно ни для расчета скоростей движения, ни для подсчета динамической плотности особей.

Видеоматериалы разбивались на кадры для использования их в обучении системы распознавания. Изображения маркировались с помощью утилиты Label Studio. Маркировка объекта представляет собой

- ограничивающую рамку с координатами двух точек: X_{\min} , Y_{\min} – верхней левой, X_{\max} , Y_{\max} – правой нижней;

- две ключевые точки: X_a , Y_a – координаты точки, обозначающей центр брюшка, X_h , Y_h – координаты точки, обозначающей центр головы.



Всего было размечено 600 изображений, 480 из них использовались для обучения, а 120 для валидации (<https://www.kaggle.com/datasets/elizamoscovskaya/ant-2-keypoints-dataset>). Размер изображения муравья в кадре в среднем составил 32×42 пикселя. Количество полностью видимых муравьев на кадре находилось в интервале [0, 15] штук.

Распознавание

Поскольку для решения поставленной задачи требуется определить не только область объекта, но и ориентацию, заданную с помощью ключевых точек, решено было воспользоваться моделью поиска ключевых точек Keypoint-RCNN из фреймворка Pytorch. Данная архитектура построена поверх предобученной магистрали ResNet-50.

Изображение на входе модели имеет размер 224×224 пикселя, и, если сжимать имеющиеся

изображения до такого формата, муравьи на них будут практически неразличимы. Поэтому во время обработки изображение разрезается на несколько частей с небольшими перекрытиями. Для получившегося набора данных изображение разрезалось на двенадцать частей. После разбиения оригинального изображения на части каждое новое изображение сжималось до размера входа модели и аннотации нормировались на новый диапазон. Каждая часть последовательно передается на вход сети. Предсказанные значения для каждой части объединяются, а в пересекающихся зонах предварительно проходят через фильтр немаксимального подавления (non-maximum suppression). Финальные предсказания визуализируются на исходном кадре.

Точность определения ограничивающих рамок муравьев после обучения составила 0.92 по критерию mAP (mean average precision), среднеквадратическое отклонение определения ориентации составило 0.4 радиана.

Отслеживание

Как уже упоминалось, для задач, связанных с изучением поведения муравьев, недостаточно уметь определять их на отдельных кадрах – требуется сопоставлять кадры между собой для построения траекторий, хранящих историю перемещения особи, хронометраж движения и т.п. Однако ни одна существующая на сегодняшний день архитектура нейронной сети не дает стопроцентную точность, что делает невозможным использование простого сопоставления распознанных объектов между кадрами. Поэтому для обработки некоторых ошибок распознавания предлагается использовать методы оценки положения.

Основной проблемой отслеживания является пересечение траекторий нескольких объектов. Происходит это из-за того, что в момент наложения объектов система распознавания иногда определяет не каждый объект в отдельности, а один слитный. Такое распознавание может привести к потере траектории одного из объектов. То есть после разъединения объектов отслеживание одного из них начнется только с этого момента, а предыдущая история будет утеряна. Кроме того, возможны случаи, когда обе траектории будут сохранены, однако после потери объектов во время пересечения сопоставятся неверно. Во избежание этого нужно реализовать метод отслеживания таким образом, чтобы его прогноз строился на основе не только текущего положения объекта, но и его ориентации, которая, в частности, поможет корректно оценить скорость движения объекта. Предполагается, что такая информация позволит не потерять ни один из объектов в момент наложения.

Отслеживание выполняется с использованием расширенного фильтра Калмана, который предсказывает новое состояние по предыдущему и в последующем обновляет это состояние по внешним измерениям. Для работы фильтра необходимо определить переменные, которые будут описывать состояние системы, а также уравнения, выражающие их изменение во времени.

Вектор наблюдения (значения, получаемые от системы детектирования) имеет следующий вид:

$$Z = [x, y, \alpha], \quad (1)$$

где x, y – координаты центра ограничивающей рамки объекта; α – угол между осью OX и вектором, проходящим через распознанные ключевые точки брюшка и головы.

Вектор состояния описывает состояние объекта пятью характеристиками:

$$X = [x, y, \alpha, v, w], \quad (2)$$

где первые три значения совпадают с Z (1); v – линейная скорость объекта; w – угловая скорость объекта.

Порядок ковариационной матрицы шума измерений (R) равен длине вектора наблюдения Z (1). Принято допущение, что измерения не коррелируют между собой, поэтому матрица имеет диагональный вид, где значения – дисперсия для каждого измеряемого параметра, рассчитанная на тестовой выборке полученного набора данных. Были получены значения $\sigma_x^2 = 1,22$, $\sigma_y^2 = 1,75$, $\sigma_\alpha^2 = 0,39$.

Предсказание нового состояния X (2) на каждом шаге осуществляется с помощью функций:

$$f_1(X_{t_0}, dt) = x_{t_0+dt} = X_x + X_v \times \cos X_\alpha \times dt,$$

$$f_2(X_{t_0}, dt) = y_{t_0+dt} = X_y + X_v \times \sin X_\alpha \times dt,$$

$$f_3(X_{t_0}, dt) = \alpha_{t_0+dt} = X_\alpha + X_w \times dt,$$

$$f_4(X_{t_0}, dt) = v_{t_0+dt} = X_v \times k^N,$$

$$f_5(X_{t_0}, dt) = w_{t_0+dt} = X_w \times k^N,$$

где параметр N определяет количество идущих подряд шагов, на которых не происходило сопоставление предсказанного фильтром значения с данными, полученными от блока распознавания; k – коэффициент затухания скорости. Отсутствие такой поправки, то есть принятие $k = 1$, грозит потерей траекторий ввиду разрыва из-за пересечения объектов либо резкой остановки, в то время как при $k < 1$ (в данной работе $k = 0.9$) у фильтра появляется возможность адаптироваться во время резкой остановки объекта, которая присуща исследуемым объектам. На рисунке 2а показана работа фильтра при $k = 1$. Во время наложения объектов друг на друга траектория одного из муравьев продолжает двигаться в сторону его последней предсказанной ориентации, из-за чего впоследствии связывается с муравьем, находящимся справа, а отслеживание муравья из пересечения начинается заново. На рисунке 2б показана работа расширенного фильтра Калмана при $k = 0.9$. За счет такой поправки траектория ожидает появления объекта в области наложения, поэтому после разъединения объектов отслеживание продолжается верно и без потери истории перемещения.

Предсказанное положение рассчитывается фильтром исходя из информации о предыдущем положении. Новое положение для каждого фильтра назначается в момент обновления

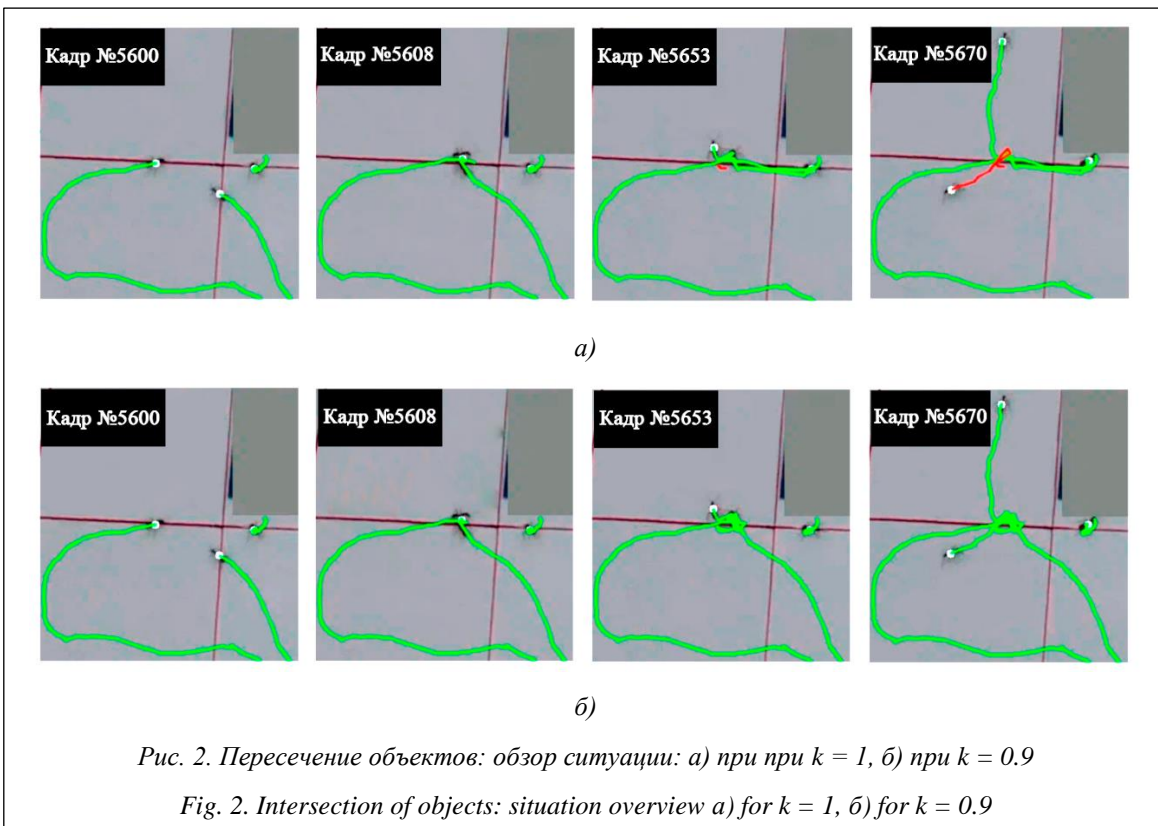


Рис. 2. Пересечение объектов: обзор ситуации: а) при $k = 1$, б) при $k = 0.9$

Fig. 2. Intersection of objects: situation overview а) for $k = 1$, б) for $k = 0.9$

предсказанных значений по внешним измерениям. Для набора существующих траекторий X из (2) и полученных для текущего кадра новых положений Z из (1) строится матрица времени перемещений по формуле

$$t(Z, X) = \max \left(\frac{\sqrt{\sum_{i=1}^{n=2} (X_i - Z_i)^2}}{X_v}, \frac{X_\alpha - Z_\alpha}{X_w} \right).$$

Затем жадным алгоритмом из матрицы изымаются значения, соответствующие минимальному времени транспортировки, пока не превысят некоторого порога, который был выбран равным 2 секундам. Выбранные таким образом Z используются для обновления соответствующих им X . Оставшиеся значения Z начинают новые траектории. Значит, у некоторого ложного срабатывания есть шанс породить ложную траекторию, которая может, в свою очередь, негативно повлиять на результат аналитической системы.

Чтобы не отслеживать ложные срабатывания на протяжении всего видеофайла, предложено добавить каждой траектории состояние, описывающее степень уверенности в ее достоверности. Это позволяет отсеять траектории, строящиеся для ложных срабатываний детектора, на раннем этапе. Состояние может принимать одно из трех значений.

Неподтвержденное. По умолчанию присваивается всем новым траекториям. Такие траектории будут обновляться ограниченное число раз. Если за это время траектория не перейдет в состояние подтвержденной, она удаляется и обновляться больше не будет.

Подтвержденное. Траектории с таким состоянием считаются истинными и будут участвовать в дальнейшем анализе.

Сохраненное. Это состояние присваивается подтвержденным траекториям, которые не обновлялись определенное количество шагов. Обычно такое состояние присваивается муравьям, которые вышли за границы кадра или исчезли из поля зрения камеры по другим причинам.

По завершении анализа расширенным фильтром Калмана формируется список с подтвержденными и сохраненными траекториями вместе с временной меткой, представляющей собой номер кадра, с которого началось отслеживание объекта.

Аналитическая система

Для подсчета динамической плотности выбирается некоторая область. Затем полученные от системы отслеживания траектории разбиваются на интервалы по одной минуте и координаты

наты каждого интервала проверяются на принадлежность интересующей области.

Если муравей побывал в выделенной области, а затем покинул ее, то при повторном заходе в область этот же муравей считается новым. Это сделано для упрощения сравнения данных, полученных при обработке видеозаписи системой и при обработке специалистом вручную. При расчете динамической плотности вручную мирмеколог не отслеживает траектории муравьев и каждый повторный заход муравья на площадку расценивается как новый. Однако это не распространяется на краткосрочные выходы из области. Чтобы отсеивать данные ситуации, была внедрена эвристика, которая определяет такие участки траектории и игнорирует их.

Эксперименты

Для проверки качества функционирования аналитической системы была разработана методика эксперимента с учетом того, что динамическая плотность, рассчитанная человеком, также может содержать ошибки. Для их выявления расчет динамической плотности осуществляется четырьмя способами:

- первый подсчет человеком вручную;
- автоматический подсчет;
- повторный подсчет человеком после нанесения системой на видеозапись траектории движения муравьев;
- рассмотрение человеком всех расхождений результатов, полученных разными методами, и определение места возникновения ошибки и ее причины.

Результаты

Работа системы проверялась на двух видеозаписях, выполненных в полевых условиях на искусственной размеченной поверхности (https://github.com/moskovskayaliza2002/ant_detection). На обеих видеозаписях наблюдение за объектами производилось сверху, что соответствует поставленным условиям. Частота кадров каждой видеозаписи составила 30 кадров в секунду. Продолжительность первой видеозаписи (<https://youtu.be/an5K5Kt5Sws>) – 21 мин. 27 сек., разрешение – 1 520×1 080 пикселей. Продолжительность второй видеозаписи (<https://youtu.be/Agx6eAVCC3w>) – 10 мин. 6 сек., разрешение – 1 920×1 080 пикселей. Отличие данной видеозаписи от предыдущей за-

ключается в перекрытии ее центра (<http://www.swsys.ru/uploaded/image/2024-1/6.jpg>), где находится кормушка. Кормушка является местом скопления муравьев и не участвует в расчете плотности. Такое видео позволит проверить работу системы с невидимыми участками, поскольку в условиях наблюдения за объектами в реальной среде такие ситуации вероятны. Для обеих видеозаписей была выбрана область площадью 1 дм², в которой производился подсчет динамической плотности. Временной интервал для определения плотности составлял 1 минуту.

Результаты подсчета динамической плотности каждым способом для первой видеозаписи представлены на рисунке 3, для второй – на рисунке 4.

Поскольку четвертый способ оценки динамической плотности был выбран как истинный, точность остальных рассчитывалась в сравнении с ним. Результаты точности расчета динамической плотности представлены в таблице 2. Точность рассчитывается по формуле

$$accuracy = \left(1 - \frac{\sum |e|}{gt} \right) \times 100 \%,$$

где e – число муравьев, на которое ошибся способ за один временной интервал; gt – сумма истинных значений плотности за все время.

Таблица 2

Точность способов подсчета на видеозаписях

Table 2

Accuracy of calculating methods in video recordings

Видео	Метод		
	Ручной	Ручной с треками	Автоматический
№ 1	92,45 %	96,23 %	92,45 %
№ 2	89,55 %	95,52 %	91,04 %

Сравнивая методы с точки зрения времени, затраченного на подсчет данной характеристики, можно увидеть результаты, отраженные в таблице 3.

Автоматический расчет производился на компьютере с видеокартой nVidia GeForce RTX 2060 SUPER и процессором Intel(R) Core(TM) i5-9400F. Время обработки видеозаписи системой на данный момент выше, чем мирмекологом вручную. Тем не менее применение машино-часов обходится значительно дешевле и позволяет специалисту использовать время для решения других задач.

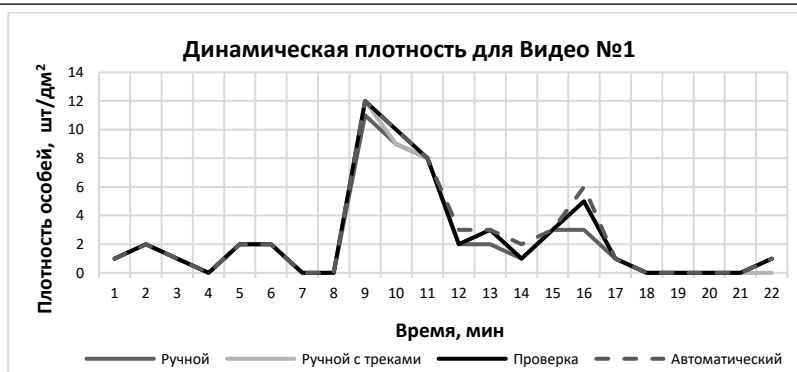


Рис. 3. Сравнение методов подсчета динамической плотности на видеозаписи № 1

Fig. 3. Comparing methods for calculating dynamic density in video recording No. 1

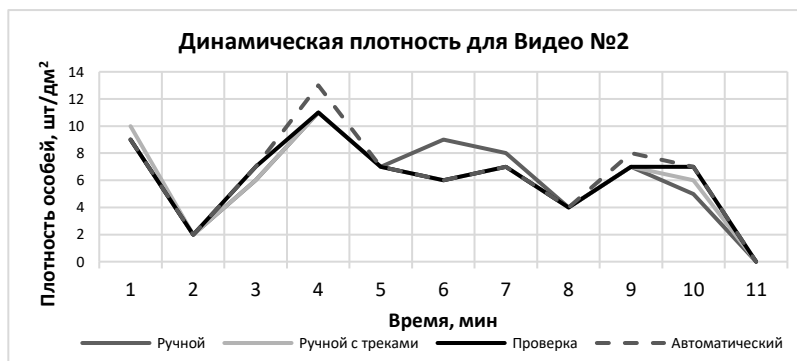


Рис. 4. Сравнение методов подсчета динамической плотности на видеозаписи № 2 с перекрытым центром

Fig. 4. Comparing methods for calculating dynamic density in video recording No. 2 with center blocked

Обсуждение результатов

Сравнение получившихся значений (табл. 2) показывает, что точность автоматического подсчета превышает точность ручного, следовательно, система уже может заменить человека как для получения более точных результатов, так и для сокращения трудоемкой и рутинной работы. Можно заметить, что точность подсчетов вручную по видеозаписям с прорисованными траекториями превышает точность всех остальных методов, что также указывает на альтернативный способ использования системы в качестве предобработчика сырых видеозаписей для дальнейшего анализа вручную и повышения эффективности мирмекологических исследований.

По результатам финальной кросс-проверки удалось выявить список ситуаций, которые вызывают неустраняемые ошибки автоматического подсчета.

Дрожание камеры во время съемки материала в полевых условиях неизбежно и ведет к

ошибкам детектора из-за смазанных кадров и к смещению границ выделенной для подсчета динамической плотности области, что также может повлиять на подсчет.

Отсутствие четкого определения наличия муравья в области. При подсчете человеком вручную нет однозначного критерия, определяющего заход муравья в область подсчета. Критерии оценки разнятся у разных людей. Более того, из-за характера движения муравьев и ракурса камеры пересечения границ воспринимаются одним и тем же человеком от просмотра к просмотру по-разному. Все эти факторы приводят к небольшой разнице результатов экспертов. В силу этого получаются расхождения в подсчетах, сделанных человеком и произведенных автоматически, поскольку в отличие от человека у системы однозначный критерий. В системе отслеживается передвижение центра муравья, и, поскольку во время движения вдоль границы центр муравья может неоднократно пересекать эту границу, плотность искусственно увеличивается. Введенная эвристика, кото-

Таблица 3

Время для обработки видеозаписей

Table 3

Time for processing video recordings

Видео	Продолжительность видеозаписи	Метод		
		Ручной (человеко-мин.)	Ручной с треками (без учета времени предобработки системой) (человеко-мин.)	Автоматический (машино-мин.)
№ 1	21 мин. 27 сек.	67 мин.	40 мин.	310 мин.
№ 2	10 мин. 6 сек.	61 мин.	52 мин.	154 мин.

рая находит кратковременные участки выхода муравья из отслеживаемой области и игнорирует их, справляется с большинством подобных случаев. Однако факторы проблемы четкого определения иногда все-таки приводят к ошибкам.

Кратковременное появление муравья в поле зрения видеокamеры. Траектория должна успешно отслеживаться в течение определенного времени, прежде чем перейти в подтвержденное состояние, и учитываться в расчете количественных характеристик для того, чтобы отсеять ложные срабатывания детектора на других насекомых, которые неизбежно появляются в условиях полевого наблюдения. Из-за этого муравьи, появляющиеся на короткий срок в поле зрения камеры, не успевают перейти в подтвержденное состояние и не учитываются при подсчете количественных характеристик. Ошибка такого рода может возникнуть при подсчете динамической плотности в области, близко расположенной к перекрытиям.

При ручном подсчете появляются ошибки, связанные со сложностью отслеживания муравьев с высокой скоростью передвижения и с одновременным отслеживанием более двух особей, пересекающих границы области подсчета. Такие ситуации приводят к многократному просматриванию видео на сниженной скорости воспроизведения, что видно из таблицы 3. В случае анализа видео человеком скорость обработки зависит не столько от длины самого видео, сколько от его содержания, в то время как при автоматическом распознавании время обработки прямо пропорционально продолжительности видеозаписи.

На данном этапе разработки системы и на существующем оборудовании время обработки видеопоследовательности машиной больше времени обработки той же видеозаписи мирмекологом в 2–3 раза. Тем не менее, как уже говорилось, машино-часы дешевле человеко-

часов. Кроме того, система осуществляет трекинг муравьев и хронометраж их движения, что потенциально дает возможность параллельно решать другие задачи. Примеры таких задач – расчет скорости муравьев (средней по траектории), скорости бега (без остановок), описание последовательностей действий.

Точность определения динамической плотности особей системой выше или равна точности определения человеком вручную при просмотре видеозаписи. Во всех случаях она составила более 90 %. Сравнение с подсчетом динамической плотности в поле не проводилось, но можно обоснованно предположить, что текущий результат, демонстрируемый системой, в среднем лучше того, что может получить человек в поле. Это связано с тем, что камера не раздражает муравьев так, как человек-наблюдатель, система видеорегистрации не устает, не отводит взгляд, не моргает. Система распознавания вполне применима для сокращения времени работы мирмеколога, однако на данном этапе требуется частичная проверка истинности данных.

Можно выделить ряд направлений дальнейшего развития исследований:

- адаптация системы для распознавания муравьев на естественном субстрате (почве, подстилке, стволах деревьев и т.д.); важно, что в траве динамическую плотность особей не рассчитывают;
- расчет других параметров поведения муравьев (интенсивность движения на дорогах, скорость и т.д.);
- повышение точности работы системы, чтобы исключить или свести к минимуму необходимость проверки данных вручную.

Заключение

Разработана система автоматического определения динамической плотности муравьев на

территории на основании видеозаписей, сделанных в полевых условиях. Система состоит из нескольких частей – блоков распознавания, отслеживания и анализа. Первая часть – сверточно-нейронный детектор. Он был обучен с точностью определения ограничивающих рамок, равной 0.92 по критерию mean average precision, и со среднеквадратическим отклонением определения ориентации 0.4 радиана, что позволяет различать пересекающиеся траектории в большинстве случаев. Вторая часть выполняет отслеживание на основе расширенного фильтра Калмана с использованием дополни-

тельной фильтрации траекторий по степени уверенности. Третья часть, аналитическая система расчета динамической плотности, показала среднюю точность, равную 93 % на двух видео, что сравнимо и немного превышает точность анализа человеком вручную. Полученные результаты точности, а также временных характеристик расчета позволяют говорить о применимости системы в мирмекологических исследованиях. Дальнейшая работа авторов будет направлена на развитие системы и расширение спектра задач автоматического подсчета.

Список литературы

1. Карпов В.Э., Карпова И.П., Кулинич А.А. Социальные сообщества роботов. М.: УРСС, 2019. 356 с.
2. Бургов Е.В., Московский А.Д. Разработка системы автоматического распознавания объектов на видеозаписи для мирмекологических исследований // Муравьи и защита леса: матер. XV Всеросс. мирмекологического симпозиума. 2018. С. 32–35.
3. Ślipiński P., Żmihorski M. Changes in the speed of ants as a result of aggressive interactions. *Insect Sci.*, 2017, vol. 24, no. 5, pp. 842–852. doi: 10.1111/1744-7917.12354.
4. Захаров А.А., Длусский Г.М., Горюнов Д.Н. и др. Мониторинг муравьев Формика. М.: КМК, 2013. 99 с.
5. Захаров А.А. Муравьи лесных сообществ, их жизнь и роль в лесу. М.: КМК, 2015. 404 с.
6. Sclocco A., Ong S.J.Y., Pyau Aung S.Y., Teseo S. Integrating real-time data analysis into automatic tracking of social insects. *R. Soc. Open Sci.*, 2021, vol. 8, art. 202033. doi: 10.1098/rsos.202033.
7. Imirzian N., Zhang Y., Kurze C., Loreto R.G., Chen D.Z., Hughes D.P. Automated tracking and analysis of ant trajectories shows variation in forager exploration. *Sci. Rep.*, 2019, vol. 9, art. 13246. doi: 10.1038/s41598-019-49655-3.
8. Wu M., Cao X., Guo S. Accurate detection and tracking of ants in indoor and outdoor environments. *bioRxiv*, 2020. URL: <https://www.biorxiv.org/content/10.1101/2020.11.30.403816v1.full> (дата обращения: 15.06.2023). doi: 10.1101/2020.11.30.403816.
9. Cao X., Guo S., Lin J., Zhang W., Liao M. Online tracking of ants based on deep association metrics: Method, dataset and evaluation. *Pattern Recognition*, 2020, vol. 103, art. 107233. doi: 10.1016/j.patcog.2020.107233.
10. Wu M., Cao X., Yang M., Cao X., Guo S. A dataset of ant colonies' motion trajectories in indoor and outdoor scenes to study clustering behavior. *GigaScience*, 2022, vol. 11, art. giac096. doi: 10.1093/gigascience/giac096.

A system for recognizing, tracking and describing ant behavior on video footage taken in the field

Elizaveta D. Moskovskaya ¹✉, Evgeny V. Burgov ¹, Anton D. Moscovsky ¹, Natalia A. Grevtsova ¹

¹ Kurchatov Complex of NBICS-Nature-Like Technologies,
Federal State Budgetary Institution
"National Research Center "Kurchatov Institute",
Moscow, 123182, Russian Federation

For citation

Moskovskaya, E.D., Burgov, E.V., Moscovsky, A.D., Grevtsova, N.A. (2024) 'A system for recognizing, tracking and describing ant behavior on video footage taken in the field', *Software & Systems*, 37(1), pp. 83–94 (in Russ.). doi: 10.15827/0236-235X.142.083-094

Article info

Received: 11.07.2023

After revision: 21.08.2023

Accepted: 05.09.2023

Abstract. The work is dedicated to developing a system for recognizing, tracking, and assessing quantitative behavioral characteristics of ants in video sequences. Such software is essential for biological research, specifically for studying ants (myrmecology). The current version of the system focuses on calculating the dynamic density of specimens in a given area,

specifically the number of worker ants in a designated region per minute. Determining the dynamic density of ants in an area involves considering videos recorded on artificially marked surfaces in field conditions. Individual ants are recognized using a neural network-based detector that determines their position and key points corresponding to their head and abdomen. The object orientation in space is calculated based on these key points. The accuracy of determining ant boundary images was 92% with an orientation detection error of 0.4 radians. Due to determining object's orientation, we use an extended Kalman filter for tracking; the filter considers the tracking problem in motion prediction. The data correlation problem is solved by a criterion of minimum ant travel time. The ant movement trajectories obtained during the system operation are used to calculate dynamic density in arbitrary regions. The tracking system uses several developed heuristic techniques to reduce false-positive detections. Due to the employed algorithms, automatic calculation of dynamic density has become as accurate as manual video processing by humans. The obtained results will significantly facilitate myrmecologists' work with video recordings, replacing manhours with more cost-effective machine hours. In the future, the system can be expanded with modules for determining other quantitative characteristics of animal movement.

Keywords: detection, tracking, ants, dynamic density, convolutional neural network, extended Kalman filter

Acknowledgements. The work was financed from the state assignment of the National Research Center Kurchatov Institute

References

1. Karpov, V.E., Karpova, I.P., Kulinich, A.A. (2019) *Social Communities of Robots*. Moscow, 356 p. (in Russ.).
2. Burgov, E.V., Moscovsky, A.D. (2018) 'Development of an automatic object recognition system on video recordings for myrmecological research', *Proc. XV Myrmecological Symposium Ants and Forest Protection*, pp. 32–35 (in Russ.).
3. Ślipiński, P., Żmihorski, M. (2017) 'Changes in the speed of ants as a result of aggressive interactions', *Insect Sci.*, 24(5), pp. 842–852. doi: 10.1111/1744-7917.12354.
4. Zakharov, A.A., Dlusskiy, G.M., Goryunov, D.N. et al. (2019) *Monitoring of Formica Ants*. Moscow, 99 p. (in Russ.).
5. Zakharov, A.A. (2015) *Ants in forest communities: their life and role in the forest*. Moscow, 404 p. (in Russ.).
6. Sclocco, A., Ong, S.J.Y., Pyay Aung, S.Y., Teseo, S. (2021) 'Integrating real-time data analysis into automatic tracking of social insects', *R. Soc. Open Sci.*, 8, art. 202033. doi: 10.1098/rsos.202033.
7. Imirzian, N., Zhang, Y., Kurze, C., Loreto, R.G., Chen, D.Z., Hughes, D.P. (2019) 'Automated tracking and analysis of ant trajectories shows variation in forager exploration', *Sci. Rep.*, 9, art. 13246. doi: 10.1038/s41598-019-49655-3.
8. Wu, M., Cao, X., Guo, S. (2020) 'Accurate detection and tracking of ants in indoor and outdoor environments', *bioRxiv*, available at: <https://www.biorxiv.org/content/10.1101/2020.11.30.403816v1.full> (accessed June 15, 2023). doi: 10.1101/2020.11.30.403816.
9. Cao, X., Guo, S., Lin, J., Zhang, W., Liao, M. (2020) 'Online tracking of ants based on deep association metrics: Method, dataset and evaluation', *Pattern Recognition*, 103, art. 107233. doi: 10.1016/j.patcog.2020.107233.
10. Wu, M., Cao, X., Yang, M., Cao, X., Guo, S. (2022) 'A dataset of ant colonies' motion trajectories in indoor and outdoor scenes to study clustering behavior', *GigaScience*, 11, art. giac096. doi: 10.1093/gigascience/giac096.

Авторы

Московская Елизавета Дмитриевна¹,

лаборант-исследователь,
moscovskayaliza@gmail.com

Бургов Евгений Вадимович¹, к.б.н.,
начальник группы, burgov.ev@yandex.ru

Московский Антон Дмитриевич¹,
начальник группы, moscovskyad@gmail.com

Гревцова Наталья Андреевна¹,
инженер-исследователь, grev-natik@yandex.ru

Authors

Elizaveta D. Moskovskaya¹,

Research Assistant,
moscovskayaliza@gmail.com

Evgeny V. Burgov¹, Cand. of Sci. (Biology),
Head of Group, burgov.ev@yandex.ru

Anton D. Moscovsky¹, Head of Group,
moscovskyad@gmail.com

Natalia A. Grevtsova¹, Research Engineer,
grev-natik@yandex.ru

¹ Курчатовский комплекс
НБИКС-природоподобных технологий,
НИЦ «Курчатовский институт»,
г. Москва, 123182, Россия

¹ Kurchatov Complex
of NBICS-Nature-Like Technologies,
National Research Center "Kurchatov Institute",
Moscow, 123182, Russian Federation