

УДК 618.31.05

И. Р. Арсенюк, к. т. н., доц.; С. И. Перевозников, д. т. н, проф.;
В. В. Войтко, к. т. н., доц.; В. С. Гончарук

СИСТЕМА РАСПОЗНАВАНИЯ РАДИОУПРАВЛЯЕМОЙ ТЕЛЕЖКИ СО СТАТИЧЕСКОГО ИЗОБРАЖЕНИЯ

В статье представлены математическая модель и алгоритм распознавания радиоуправляемой тележки. На основе предложенного алгоритма разработана программа распознавания тележки. Результаты тестирования на изображениях тележки разного качества, яркости и с разной степенью помех свидетельствуют про удовлетворительную работу программы.

Ключевые слова: радиоуправляемая тележка, структурно-инвариантный алгоритм, система распознавания, нейронная сеть, полином Лагранжа.

Введение

На сегодняшний день можно наблюдать стремительное развитие робототехники и значительные достижения в области создания систем искусственного интеллекта. Поэтому актуальным является проведение исследований связанных с робототехникой, в области компьютерного распознавания разных механических объектов, в частности радиоуправляемой тележки. Такая задача, безусловно, является актуальной для многих сфер деятельности человека. Например, похожие системы активно разрабатываются в военной сфере, потому что использование радиоуправляемых объектов является намного безопаснее, чем использование человеческого вмешательства (например, проведение саперных работ). Также это направление исследований успешно развивается в отраслях, где использование человеческой работы очень опасно для здоровья (например, химическая и радиационная разведка) или вообще невозможно (космическое и глубоководное пространства с экстремальными значениями давления/температуры). Решение задачи распознавания предполагает обработку данных довольно большого объема, которые содержатся в изображении, поэтому для ее решения целесообразно использовать систему искусственного интеллекта [1].

Объектом исследования является радиоуправляемая тележка, которую необходимо распознать на статической картинке. Предмет исследования – система, способная распознавать объект в процессе нормального сценария распознавания (когда тележка присутствует на картинке полностью и качество картинки достаточное для решения задачи распознавания) и в случае возникновения проблем при распознавании (например, когда тележка присутствует на картинке частично или когда картинка очень малого размера или на ней много помех). Целью работы является разработка системы, которая эффективно решает задачу распознавания сложного объекта, а именно: радиоуправляемой тележки при условии различного качества входящего изображения.

Анализ современного состояния проблемы

В последние годы возрос интерес к исследованию и построению систем автоматического распознавания образов и машинного обучения. Примеров таких систем довольно много. Например, были получены успешные результаты при создании устройств и программ чтения напечатанных символов, в этих программах решается задача распознавания символов, обработки электрокардиограмм, распознавания языка, отпечатков пальцев и интерпретации фотоснимков. В качестве других примеров можно привести программы, распознающие символы рукописного ввода, классифицирующие сейсмическую активность, находящие вражеские объекты, выполняющие прогнозирование погоды и т. п. [1]. Приведем примеры

систем, которые могут быть использованы для решения задачи распознавания радиуправляемой тележки.

– Автоматическая классификация данных, полученных дистанционно [2].

Сравнительно недавно в Соединенных Штатах Америки возник интерес к состоянию окружающей среды и природных ресурсов, который привёл к написанию ряда приложений распознавания образов. Наибольшее внимание среди них представляет автоматическая классификация данных, полученных дистанционно. Поскольку объем данных, получаемых от многодиапазонных спектральных портативных устройств, установленных на самолетах, спутниках и космических станциях, чрезвычайно большой, возникает необходимость обратиться к автоматическим средствам обработки и анализа этой информации. Дистанционный сбор данных используется при решении разных задач. Среди отраслей, вызывающих особую заинтересованность можно отметить, например, землеиспользование, выявление заболеваний сельскохозяйственных культур, лесное хозяйство, контроль качества воздуха и воды, геологические и географические исследования, прогнозирование погоды и много других задач, связанных с охраной окружающей среды [2].

– Среда разработки систем машинного зрения National Instruments Vision Development Module (NI Vision DM) [3].

Среда NI Vision DM (версия 8.5) предназначена для инженеров и ученых, использующих технологии технического зрения в промышленных и научных задачах. Модуль содержит в себе интерактивную оболочку NI Vision Assistant, предназначенную для разработки прототипов приложений без программирования и мощную библиотеку с более чем 200 функциями обработки изображений IMAQ Vision [3].

Возможность совместного использования NI Vision Assistant и IMAQ Vision позволяет сократить сроки создания систем машинного зрения, поскольку NI Vision Assistant может осуществлять автоматическую генерацию кодов в Lab VIEW, C/C++ и Visual Basic с той же последовательностью операций, которая использовалась при работе в NI Vision Assistant. Здесь появляется возможность интегрировать диаграмму видеособирания в промышленную систему измерений и автоматизации, которая, кроме этого, может включать в себя управление движением, измерительными приборами или устройствами ввода/вывода сигналов.

Эта среда является довольно мощной и универсальной, но она имеет несколько неудобный интерфейс, в котором тяжело настроить детали работы системы и добавлять в систему собственные алгоритмы распознавания.

– Программный комплекс Ident Smart Studio предметно независимого распознавания графических объектов [4].

Еще одним мощным программным продуктом, разработанным для решения подобных задач является система распознавания сложных объектов Ident Smart Studio [4], в которой разработан программный модуль распознавания Pattern Searcher. Данная система предназначена для распознавания широкого класса объектов, в том числе биологических объектов и живых животных. Она может быть адаптирована и настроена пользователем на распознавание нужного класса объектов, что свидетельствует об универсальности использования этого программного продукта. Однако о данной системе была опубликована лишь общая информация и предоставлена возможность свободного копирования демонстрационной версии модуля распознавания Pattern Searcher 1.0 Demo [4], реализующего лишь минимум возможностей ее работы на одном примере объекта распознавания и не пригодного для распознавания широкого класса объектов.

Краткий обзор основных методов распознавания изображений

Современный этап развития вычислительной техники отмечается интенсивным поиском новых принципов обработки информации, продиктованными требованиями высокой производительности систем, а также их гибкости и надежности.

При работе с изображениями следует решить задачи, которые возникают в связи с тремя основными проблемами: описание (моделирование) изображения; разработка и выбор математических средств обработки и анализа изображений; аппаратная реализация математических методов работы с изображениями. Поэтому на современном этапе можно выделить такие основные методы распознавания:

- лингвистические методы;
- эвристические методы;
- математические методы.

Лингвистический (синтаксический) метод. Если описание образов осуществляется с помощью подобразов и их соотношений, то для конструирования системы распознавания используют лингвистический, или синтаксический, подход с использованием принципа схожести свойств. Образы, принадлежащие одному и тому же классу, имеют ряд общих свойств или признаков, которые отображают сходство таких образов. Эти общие свойства можно частично ввести в память системы распознавания. Когда системе предоставить неклассифицированный образ, то выбирается набор определенных признаков, которые потом сравниваются с признаками, заложенными в памяти системы распознавания. При использовании данного метода основная задача состоит в выделении таких общих свойств объектов распознавания для их дальнейшей классификации. Часто выбор таких свойств зависит от разработчика системы распознавания, а также от предметной отрасли, где будет использоваться данная система распознавания.

Математический метод. В основу математического метода положены правила классификации, которые формулируются и выводятся в рамках определенного математического формализма с помощью принципов всеобщности свойств и кластеризации. Когда образы определенного класса представляют собой векторы, компонентами которых являются действительные числа, то этот класс можно рассматривать как кластер. Построение системы распознавания, которая базируется на реализации данного принципа, определяется пространственным размещением отдельных кластеров. Если кластеры, которые относятся разным классам, разнесены далеко друг от друга, то можно пользоваться простыми схемами распознавания, например, классификацией по принципу минимального расстояния.

Эвристический анализ. Основа эвристического анализа – интуиция и опыт человека. В этом анализе используются принципы перечисления членов класса и их общие свойства. Часто системы, построенные по таким методам, содержат набор специфических процедур, разработанных для конкретных задач распознавания. Это означает, что структура и качество эвристической системы в значительной мере определяется навыками и опытом работы тех, кто ее разрабатывает.

Анализ методов решения поставленной задачи

При разработке системы распознавания нужно проанализировать известные методы и способы, которые можно использовать для решения этой задачи и выбрать из них наиболее рациональный. Такими являются структурно-инвариантный метод и способ распознавания на основе нейронной сети.

И так, рассмотрим структурно-инвариантный метод для построения формального описания и обработки данных, распознавания сложных изображений, где общим свойством объектов распознавания является цвет.

Структурно-инвариантные алгоритмы предназначены для построения формального описания односвязных объектов на основе анализа информации о взаимном расположении на контуре точек локального экстремума кривизны и классификации таким образом образованных описаний [5].

Алгоритм строит систему признаков исследуемого объекта на основе нахождения специальных точек – полюсов при интерполяции контура. Контур представляется с помощью параметрических функций $x(t) = x_b$, $y(t) = y_b$, $t = 1, \dots, N$, где N – количество точек

контура. Для вычисления значений полюсов используется параметрическое представление интерполяционного полинома третьей степени в форме Лагранжа. Для функций $x(t)$ и $y(t)$ этот полином имеет вид:

$$P_{iii} = -\frac{(t_2 - t_1)^2}{(t_2 - t_0)(t_1 - t_0)} P_{000} + \frac{(t_2 - t_0)^2}{(t_2 - t_1)(t_1 - t_0)} P_{111} - \frac{(t_1 - t_0)^2}{(t_2 - t_1)(t_2 - t_0)} P_{222}, \quad (1)$$

где P_{iii} – значение полюса; P_{000} , P_{111} , P_{222} – значение в узловых точках интерполяции; t_i – элементы неубывающей последовательности, сформированной из координат узловых точек.

Полюса P_{iii} располагаются близко от геометрического объекта, и произвольное линейное изменение параметра t не влияет на геометрическое расположение полюсов.

$$t = Ta + b, \quad (2)$$

где $a, b \in R$ (R – область нахождения объекта), T – период дискретизации. Такая замена означает композицию преобразования переноса и преобразование сходства. Следовательно, величины расстояний полюсов к кривым $x(t)$ и $y(t)$ можно использовать как описание объекта.

Для того, чтобы определить вид каждой части контура рассматриваются уравнения касательных, проведенных в точках локальных экстремумов (ТЛЭ). Если для проведенных касательных выполняется условие

$$|k_1 - k_2| \leq \left| k - \frac{k_1 + k_2}{2} \right|, \quad (3)$$

где k, k_1, k_2 – соответственно угловые коэффициенты прямой, проведенной через ТЛЭ, и касательных, проведенных в этих точках, то принимается, что соответствующая часть контура – отрезок прямой. Если же такое условие не выполняется, то находится точка сечения нормалей к касательным (r_1), проведенных в ТЛЭ, и расстояния от этой точки к ТЛЭ (r_2). В случае, когда для величин этих расстояний выполняется условие

$$|r_1 - r_2| \leq \frac{\max(r_1, r_2)}{|r_1 - r_2|}, \quad (4)$$

где r_1, r_2 – вычисленные величины, принимаем, что часть контура – дуга круга, в противном случае часть контура – дуга эллипса.

В каждом случае части контура ставится в соответствие символ: l – если часть контура – отрезок прямой, c – если часть контура – дуга круга, e – если часть контура – дуга эллипса.

Доказано, что такая система признаков владеет инвариантными свойствами при аффинных преобразованиях объекта [5].

Для классификации объектов значения каждой функции входящего объекта сравниваются со значениями соответствующей функции представителя класса. За критерий берется минимум среднеквадратической погрешности. Если величина погрешности принимает минимальное значение при сравнении с представителем класса, то считаем, что входной объект принадлежит к тому же классу, что и представитель. Такое применение критерия минимума среднеквадратической погрешности возможно благодаря свойству периодичности функций: функция повторяет свои значения через некоторый ненулевой период, т. е. не изменяет своего значения при добавлении к аргументу фиксированного ненулевого числа (периода).

Рассмотрим использование нейросетевых технологий для решения задачи распознавания. Большой шаг в развитии для нейрокибернетики сделал нейрофизиолог Френк Розенблат, который предложил модель распознающей машины, и назвал ее “Персептрон” (от латинского *percepto* (понимаю, познаю). Во время разработки этой машины он исходил из некоторых принятых представлений о структуре мозга и зрительного аппарата. Стремясь

воссоздать функции человеческого мозга, Френк Розенблат использовал простую модель биологического нейрона и систему связей между ними [6].

Воспринимающим устройством персептрона служит фотоэлектрическая модель сетчатки – поле рецепторов, состоящее из нескольких сотен фотосопротивлений (*S-элементов*). Каждый элемент поля рецепторов может находиться в двух состояниях – возбужденном или невозбужденном, в зависимости от того, падает или нет на соответствующее фотосопротивление контур проектированной на поле фигуры. На выходе каждого элемента появляется сигнал x_i ($i = 1, 2, \dots, n$, где n – число элементов), который равен единице, если элемент возбужден, и нулю – в противном случае (рис. 1). Следующей ступенью персептрона являются ассоциативные элементы, или *A-элементы*, каждый из которых имеет несколько входов и один выход. При подготовке персептрона к эксперименту к входам *A-элемента* подключаются выходы рецепторов, причем подключение каждого из них может быть сделано со знаком плюс или минус. Выбор рецепторов, которые подключают к данному *A-элементу*, также как и выбор знака подключения, выполняется случайно. В ходе эксперимента связь рецепторов с *A-элементами* остается неизменной. *A-элементы* выполняют алгебраическое суммирование сигналов, поступивших на их входы, и полученную сумму сравнивают с одинаковой для всех *A-элементов* величиной [4].

Однако использование нейронных сетей для задач распознавания также не всегда является оптимальным, поскольку данный способ имеет и ряд недостатков. Например, не всегда существует такая комбинация весовых коэффициентов, при которой имеющееся множество образов будет удачно распознанным; отсутствие возможности, при которой можно определить, сколько времени понадобится для обучения нейросети; склонность сети к ошибочным реакциям на входной сигнал; низкая эффективность при переобучении, когда необходимо ввести новые данные, не уничтожая информации, сохранённой раньше. Все эти недостатки являются существенными и значительно усложняют процесс построения эффективной системы распознавания.

Алгоритм распознавания, основанный на методе лингвистического распознавания, где общим признаком объектов распознавания выступает цвет, является простейшим и скоростным, тем не менее он не обеспечивает достаточной точности распознавания, поскольку имеет ряд недостатков. Одним из наибольших среди таких недостатков является то, что алгоритм может оказаться совсем неэффективным при изменении яркости картинки, или в случае появления на ней световых отблесков. Поэтому данный алгоритм в процессе распознавания предлагается применять в комбинации со структурно-инвариантным, так как это повысит точность распознавания. Следовательно, при распознавании тележки появляется возможность использовать комбинированный алгоритм, состоящий из структурно-инвариантного алгоритма, который будет выполнять распознавание формы радиоуправляемой тележки, и алгоритм, основанный на методе лингвистического распознавания. Это повысит точность и быстродействие системы распознавания, а также ее помехоустойчивость.

Общий алгоритм работы системы распознавания

Разработаем общий алгоритм работы системы распознавания. На вход алгоритма подается массив данных, в котором записаны все свойства каждой точки изображения. Далее массив последовательно обрабатывается (попиксельно: от первого до последнего; также как луч электронов пробегает по экрану телевизора), во время проверки каждой точки проверяются

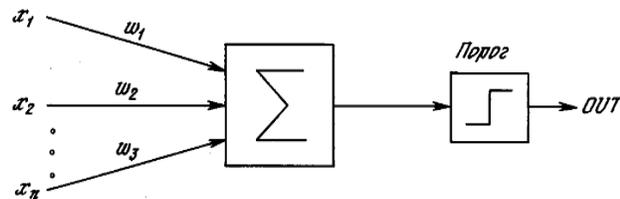


Рис. 1. Персептронный нейрон

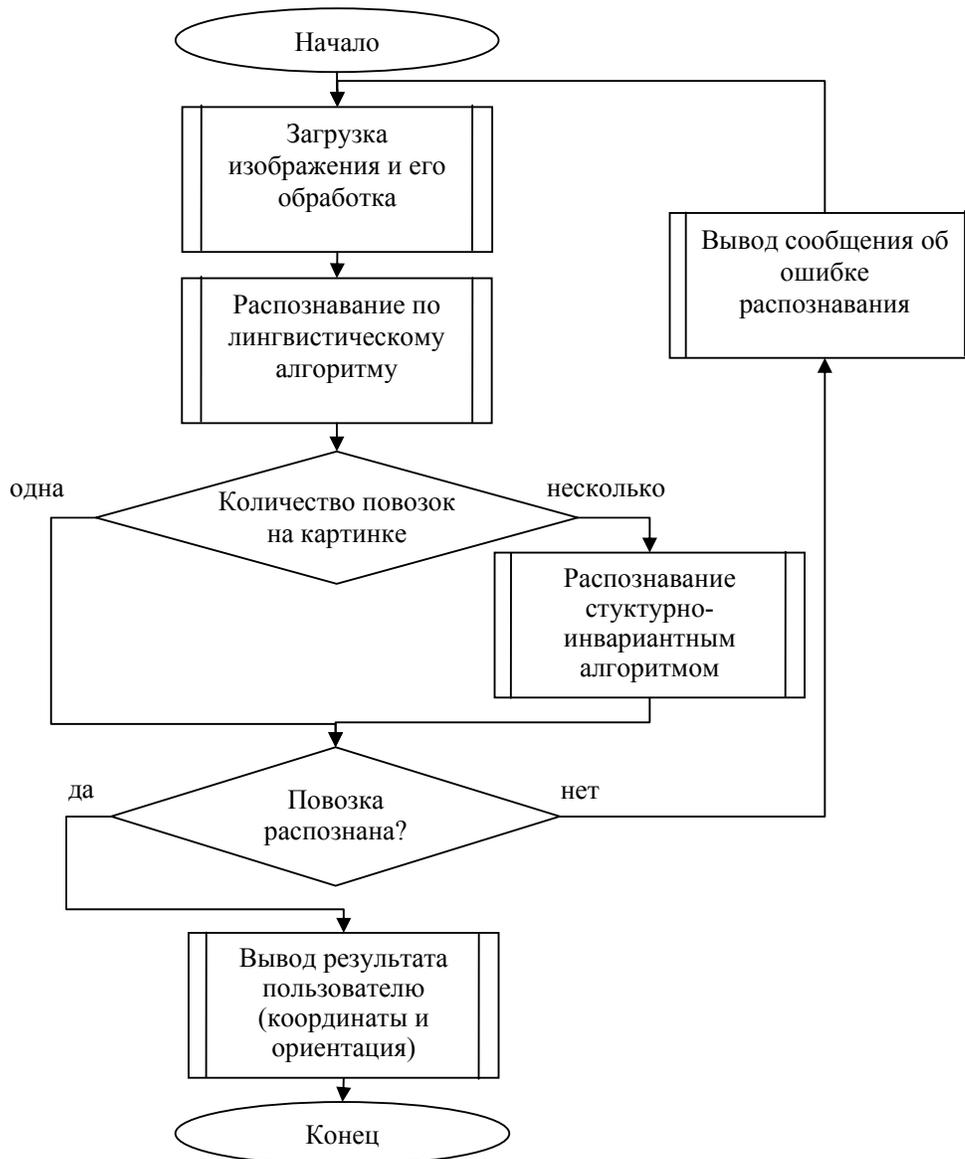


Рис. 2. Алгоритм работы основного программного модуля системы

ее свойства согласно логическим правилам алгоритма. Если точка совсем не соответствует по свойствам той, которую нужно распознать, тогда соответствующий пиксель в дальнейшем не обрабатывается и алгоритм начинает рассматривать не соседнюю точку, а точку, которая находится на определенном расстоянии от данной точки. Такой метод, как показали исследования, позволяет повысить быстродействие алгоритма до 40%. Если найдена точка, соответствующая по свойствам той, которую нужно распознать, то алгоритм начинает просматривать соседние точки, и если их свойства удовлетворяют условие распознавания, то данная область считается распознанной (например, распознана передняя часть тележки), и алгоритм продолжает свою работу до тех пор, пока не будут найдены координаты передней и задней части тележки, а также пока не будет определена его ориентация на картинке. Если алгоритм, основанный на методе лингвистического распознавания, распознал тележку не полностью или появились дополнительные сложности при распознавании (например, на картинке присутствуют несколько тележек), то выполняется алгоритм, который работает по структурно-инвариантному методу. Данный алгоритм работает несколько медленнее, но он уже работает не со всем картинке, а с данными, полученными при работе лингвистического алгоритма. Результатом работы алгоритма является координаты местонахождения тележки и

ее ориентация или сообщение о невозможности распознавания тележки.

Учитывая вышеуказанные соображения относительно основных принципов работы алгоритма распознавания радиоуправляемой тележки, получим схему, приведенную на рис. 2.

Анализ результатов тестирования программы

На основе предложенного выше алгоритма распознавания разработан программный продукт, который довольно эффективно решает задачу распознавания радиоуправляемой тележки на статической картинке. Результаты работы программы приведены ниже.

Было протестировано 30 изображений, из которых лишь 3 не были распознаны. Приводить их все нет смысла, поскольку среди них много аналогичных, поэтому приведем лишь три примера изображений, которые свойственны таким классам изображений, как качественные, яркие и некачественные (очень темные).

Так, на рис. 3 показано, что объект распознавания виден четко, и поэтому он был распознан правильно и быстро (время распознавания составило менее 40 мс, координаты: (686; 762); (1124; 1100), ориентация – 48°).

В случае чрезмерной яркости изображения и наличия на объекте большого количества отблесков, лингвистический алгоритм несколько неправильно распознал одну из координат, поэтому здесь еще применялся структурно-инвариантный алгоритм, в результате чего быстродействие системы уменьшилось, и тележка была распознана уже за 150 мс (рис. 4), координаты: (1156; 413); (1218; 885), ориентация 2° .

В случае же когда картинка очень некачественная и тележки почти не видно (рис. 5), программа анализировала картинку на протяжении 100 мс, но, не найдя объект распознавания, прекратила работу и вывела сообщение об ошибке. Для того, чтобы распознавать такие некачественные картинки, нужно их обрабатывать специальными программными фильтрами перед распознаванием, но тогда время распознавания значительно возрастает.



Рис. 3. Вид окна программы с результатом распознавания

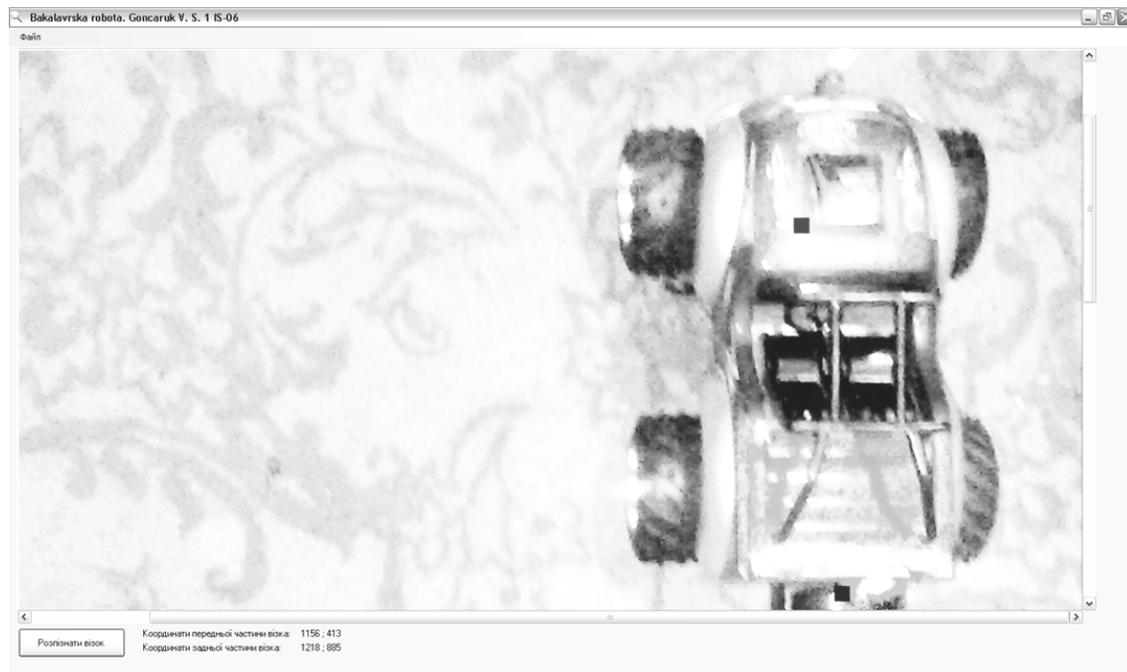


Рис. 4. Вид окна программы при неточном распознавании



Рис. 5. Вид окна программы в случае невозможности распознать тележку

Выводы

В статье исследованы основные подходы относительно разработки системы распознавания радиоуправляемой тележки на статическом изображении. Осуществлено сравнение различных методов распознавания, которое показало целесообразность использования комбинированного алгоритма распознавания, использующего лингвистический и структурно-инвариантный методы, что позволяет повысить быстродействие и точность распознавания.

Предложена схема общего алгоритма распознавания радиоуправляемой тележки. Основной недостаток данной системы – недостаточная универсальность распознавания, которая может быть улучшена при применении более универсального метода распознавания. Это позволит расширить работоспособность системы и в других предметных отраслях, а также уменьшит привязанность программы к конкретному объекту распознавания.

Разработан программный продукт, который распознает радиоуправляемую тележку со статического изображения. При испытании работы программы были протестированы тридцать рисунков тележки, из которых правильно распознаны двадцать семь. Три

нераспознанных картинки были очень некачественные и содержали много помех, которые привели к тому, что алгоритм распознавания оказался неэффективным в данных случаях. Во всех других случаях программа корректно и быстро распознала радиоуправляемую тележку.

В дальнейшем предусматривается расширение возможностей разработанного программного комплекса в плане распознавания движущихся изображений в режиме реального времени. Это позволит использовать данную программу как модуль к комплексной системе управления радиоуправляемыми объектами в реальном режиме времени.

СПИСОК ЛИТЕРАТУРЫ

1. Люгер Д. Ф. Искусственный интеллект. Стратегии и методы решения сложных проблем/ Д. Ф. Люгер – М.: Вильямс, 2005. – 864 с.
2. Ту Дж. Принципы распознавания образов / Ту Дж., Гонсалес Р. – М.: Мир, 1988. – 401 с.
3. National Instruments Announces NI Vision Development Module 8.5 Инструментарий для создания систем машинного зрения [Электронный ресурс] / National Instruments// Режим доступа: <http://industrial-embedded.com/national-module-8-5>
4. Проект Ident Smart Studio, материалы [Электронный ресурс] / Мальчевский С. А. // Режим доступа: http://iss.norcity.ru/download_get_file.php?id=17
5. Косаревич Р. Я. Структурно-інваріантні алгоритми опису та розпізнавання зображень: дис. канд. техн. наук: 05.13.06 / Косаревич Р. Я. – Львів, 1999. – 150 с.
6. Круглов В. В. Искусственные нейронные сети. Теория и практика / В. В. Круглов, В. В. Борисов. – Горячая линия-Телеком, 2002. – 382 с.

Арсенюк Игорь Ростиславович – к. т. н., доцент кафедры компьютерных наук.

Первозников Сергей Иванович – д. т. н., профессор, заведующий кафедрой компьютерных наук.

Войтко Виктория Владимировна – к. т. н., доцент кафедры программного обеспечения, декан факультета компьютерного интеллекта.

Гончарук Владислав Сергеевич – студент кафедры компьютерных наук.
Винницкий национальный технический университет.