УДК 004.89

МЕТОД РАСПОЗНАВАНИЯ ОБЪЕКТОВ НА ОСНОВЕ КОМПЛЕКСНОЙ ОБРАБОТКИ ДАННЫХ СИСТЕМЫ ТЕХНИЧЕСКОГО ЗРЕНИЯ НАЗЕМНОГО БЕСПИЛОТНОГО ТРАНСПОРТНОГО СРЕДСТВА

Безумнов Д.Н., *Московский технический университет связи и информатики (МТУСИ), старший преподаватель; e-mail: d.n.bezumnov@mtuci.ru*
Чиров Д.С., *МТУСИ, заведующий кафедрой, профессор, д.т.н.; e-mail: d.s.chirov@mtuci.ru*

AN OBJECT CLASSIFICATION ALGORITHM BASED ON COMPLEX DATA PROCESSING OF THE GROUND-BASED ROBOTIC COMPLEX VISION SYSTEM

Bezumnov D.N., Chirov D.S.

The article presents the results of a study on improving the efficiency of vision systems of ground-based robotic complexes. It is proposed to use the integration of optical and radar channel information at the solution level to increase the probability of correct classification of objects in the field of view of the vision system. Classification (recognition) of objects is performed in each of the channels independently, using a convolutional neural network. Obtaining a comprehensive assessment of the classification of objects is carried out according to the Shortliff scheme. The simulation results showed that the proposed algorithm makes it possible to increase the efficiency of the vision system of the robotic complex by 43 % compared with using only the optical channel and by 38 % compared with using only the radar channel.

Key words: robotics complex, vision system, object classification, convolutional neural network, optical and radar images.

Ключевые слова: беспилотное транспортное средство, робототехнический комплекс, система технического зрения, классификация объектов, сверточная нейронная сеть, оптические и радиолокационные изображения.

В статье представлены результаты исследования по повышению эффективности систем технического зрения (СТЗ) наземных беспилотных транспортных средств (БТС). Для повышения вероятности правильной классификации (распознавания) объектов в поле зрения системы технического зрения предлагается использовать комплексирование информации от оптического и радиолокационного канала на уровне решений. Распознавание объектов производится в каждом из каналов независимо, с использованием сверточной нейронной сети. Получение комплексной оценки классификации объекта производится по схеме Шортлиффа. Результаты моделирования показали, что предложенный алгоритм позволяет повысить эффективность СТЗ наземного БТС по сравнению с системами, использующими только оптический или только радиолокационный канал.

Введение

Робототехника является одной из ключевых сквозных технологий цифровой экономики [1]. Беспилотные транспортные средства (БТС) уже активно применяются в различных отраслях экономики [2–5]: доставка грузов, сельское хозяйство, космические исследования и т.д. Все это требует совершенствования технологий робототехники и одним из таких направлений является внедрение в робототехнические комплексы (РТК) искусственного интеллекта (ИИ). Наибольшие успехи по внедрению ИИ наблюдаются в системах технического зрения (СТЗ) РТК [5, 6]. Анализ существующих наземных БРС показывает, что большинство СТЗ являются телевизионными системами (оптического и тепловизионного диапазонов) [5, 6], где в качестве элемента ИИ используются искусственные нейронные сети (ИНС) для распознавания и классификации объектов в поле зрения СТЗ.

Однако во многих случаях телевизионной и даже стереотелевизионной информации оказывается недостаточно для эффективного анализа и оценки окружающей обстановки. Одним из способов повышения информативности СТЗ является применение радиолокационных средств [6, 7]. Радиолокационные методы сейчас широко используются для сбора и анализа информации в таких областях, как обнаружение и определение координат физических объектов и информационных датчиков в заданных зонах наблюдения, а также для классификации поверхностей и физических объектов.

В [6] рассмотрена возможность комплексирования оптической и радиолокационной информации в СТЗ наземного БТС с использованием схемы Шортлиффа. Данная статья является дальнейшим развитием данных исследований.

Уровни объединения

В зависимости от уровня объединения и абстракции объединенных данных, уровни мультимодального объединения можно разделить на уровень данных, уровень признаков и уровень принятия решений. В контексте объединения радара и камеры объединение на уровне данных часто называют ранним объединением, объединение на уровне принятия решений также называют объединением на уровне объектов или поздним объединением, а объединение на уровне признаков также известно, как среднее объединение. Три основных уровня объединения имеют следующие характеристики:

• Раннее объединение: объединение выполняется на низком уровне, где исходные данные или данные пред-

варительной обработки из различных модальностей объединяются на ранней стадии. Затем интегрированные данные подвергаются дальнейшей обработке в соответствии с требованиями задачи для получения окончательного результата.

• Среднее объединение: объединение происходит на промежуточном уровне, где признаки сначала извлекаются по отдельности из каждой модальности. Затем эти признаки объединяются для создания новых тензоров признаков для последующей обработки данных.

• Позднее объединение: объединение происходит на высоком уровне, где каждая модальность обрабатывается отдельно для извлечения признаков и составления независимых прогнозов. Наконец, эти прогнозы интегрируются для получения единого результата.

Выбор уровня объединения зависит не только от требований к точности, но и от множества факторов, таких как гибкость модели и вычислительные затраты. Таблица 1 кратко описывает наиболее заметные преимущества и недостатки трех вышеупомянутых уровней объединения [8]. Что касается объединения радара и камеры, то на ранних этапах работы основное внимание уделялось позднему объединению. В последнее время, с развитием технологии глубокого обучения (deep learning (DL)) и появлением общедоступных наборов радиолокационных данных, большинство основных методов объединения радара и камеры направлены на изучение стратегий среднего объединения [8].

При раннем объединении входные данные с радаров и раннего и позднего объединений, что делает его подходящим для сетей глубокого обучения, специализирующихся на извлечении признаков. Как правило, разнородные модальные данные в объединении радара и камеры вносят разный вклад в результаты восприятия. Одна модальность может играть доминирующую роль, в то время как другая модальность предоставляет вспомогательную дополнительную информацию. Объединение на уровне признаков идеально подходит для изучения этих взаимодополняющих взаимодействий между разнородными модальностями. В рамках методов среднего объединения существует особая парадигма, известная как объединение на гибридном уровне. Этот подход усиливает информацию с различных уровней признаков, эффективно сохраняя признаки каждого этапа. Тем не менее, объединение на гибридном уровне часто не поддается интерпретации и требует экспериментов для определения оптимальной архитектуры сети на основе эмпирических знаний. Более того, модели, использующие объединение на гибридном уровне, обычно включают больше ветвей в нейронных сетях, что в некоторой степени приводит к увеличению вычислительных затрат.

В работе применяется позднее объединение, т.е. объединение на уровне принятия решений (рисунок 1).

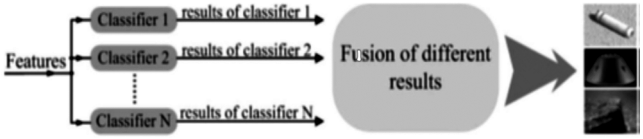


Рисунок 1. Схема объединения результатов на уровне принятия решений [9]

Метод распознавания объектов

Одной из важных задач, решаемых СТЗ, является классификация (распознавания) объектов в поле зрения. Пусть в рассматриваемой СТЗ наземного БТС имеется N информационных каналов, обеспечивающих получение оптических и радиолокационных изображений. Классификация объектов осуществляется отдельно по каждому каналу. Здесь возникает задача объединения информации на основе решений, т.е. проблема определения степени истинности заключения на основании решений, принятых отдельно по данным от двух используемых каналов.

С учётом проведённого анализа эффективным способом объединения решений является схема Шортлиффа [8]:

Таблица 1. Преимущества и недостатки различных уровней объединения

Уровень объединения	Преимущества	Недостатки
Раннее объединение	• Включает в себя наиболее полную модальную информацию	• Проблемы с большим объемом данных, значительными вычислительными затратами, чувствительностью к неправильной регистрации датчиков, низкой надежностью и ограниченной гибкостью
Среднее объединение	• Позволяет сжимать данные для повышения производительности в режиме реального времени. • Применяется для адаптивного использования синергии между разнородными модальностями на основе платформы, управляемой DL. • Обладает умеренной надежностью и гибкостью модели.	• Технология все еще находится на стадии становления, особенно в области объединения радара и камеры
Позднее объединение	• Достигается гибкая модульная конструкция и низкие затраты на связь.	• Сталкивается с непреодолимыми трудностями, связанными с потерей значительного объема ценной информации при изучении взаимодополняющих взаимодействий между разнородными модальностями

$$P_{\Sigma} = 1 - \prod_{i=1}^{N} (1 - P(A_i)), \quad (1)$$

где: N – количество измерительных каналов, Ai – информация от i-го канала, $P(Ai)$ – вероятность корректного распознавания объекта по Ai.

Алгоритм классификации (распознавания) объектов на основе информации Ai ($i \in [1, N]$), получаемой через N каналов, будет иметь вид, представленный на рисунке 2.

В случае наличия только двух каналов – оптического и радиолокационного, – алгоритм будет иметь вид, представленный на рисунке 3 [6].

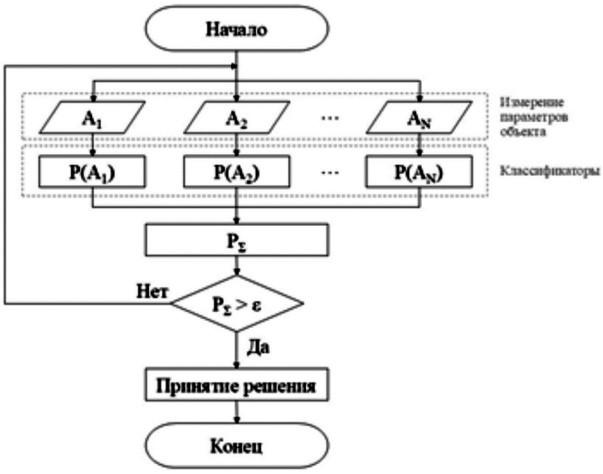


Рисунок 2. Метод распознавания объектов на основе комплексной обработки данных

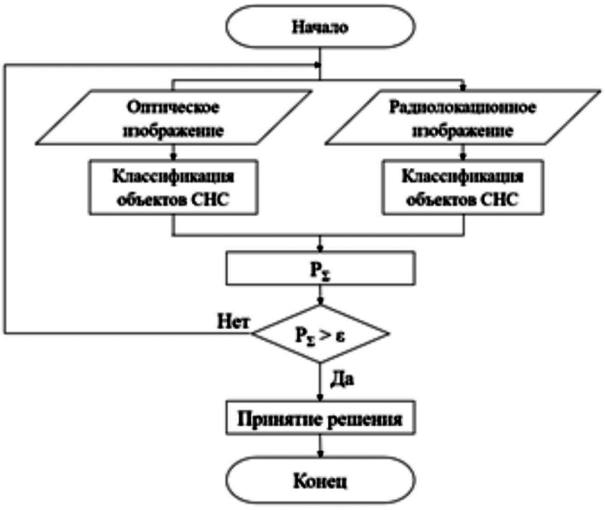


Рисунок 3. Алгоритм классификации объектов на основе комплексной обработки данных

Оценка эффективности разработанного метода

Для оценки эффективности разработанного алгоритма создана база данных оптических и радиолокационныхизображений различных объектов.

Информация об объекте, получаемая БТС через оптический канал, представляет собой RGB-изображение размером 32x32 пиксела. Информация об объекте, получаемая БТС через радиолокационный канал, представляет собой двоичное изображение размером 16x16 пикселов.

Распознавание объектов осуществляется посредством предварительно обученной СНС, модель которой представлена на рисунке 4.

СНС реализована средствами фреймворка PyTorch, предобработка данных и проведение экспериментов выполнялось при помощи фреймворков Numpy, Pandas и Scikit-learn. Нейронная сеть (см. рисунок 4) состоит из двух сверточных слоев Conv с функцией активации ReLU, слоя уменьшения размерности MaxPool, слоя выравнивания (построчное развертывание матрицы в вектор) Flatten и трех полносвязных слоев Linear, два из которых также используют функцию активации ReLU.

Оптические изображения хранятся в базе данных (БД) в виде трёхмерных массивов $\mathbf{A_i}$ с размерностью 32x32x3, значения которых являются целыми числами в диапазоне от 0 до 255. Радиолокационные изображения хранятся в БД в виде двумерных массивов $\mathbf{B_i}$ с размерностью 16x16, значения которых являются целыми числами в диапазоне от 0 до 1.



Рисунок 4. Архитектура СНС для распознавания объектов

В целях аугментации данных, т.е. искусственного увеличения объёма обучающего набора, к оптическим и радиолокационным изображениям на входе СНС применяется с вероятностью 0,5 зеркальное отображение по горизонтали.

Примеры изображений приведены на рисунке 5. БД изображений состоит из 500 оптических и 500 радиолокационных изображений для пяти классов объектов (автомобиль, птица, кошка, собака, лошадь).

При подаче на вход СНС массивы $\mathbf{A_i}$ и $\mathbf{B_i}$ преобразуются в тензоры с размерностью 3x32x32 и 1x16x16 соответственно. Значения массивов $\mathbf{A_i}$ также проходят нормировку к диапазону значений [0, 1].



а) *б)*

Рисунок 5. Примеры оптических (а) и радиолокационных изображений (б)

Были проведены эксперименты по обучению СНС распознаванию и классификации объектов на оптических и радиолокационных изображениях с параметрами, описанными в таблице 2.

Таблица 2. Параметры экспериментов по обучению СНС

Параметр	**Эксперимент №1**	**Эксперимент №2**
Канал получения изображения	Оптический	радиолокационный
Количество классов	5	5
Объём обучающего набора (количество изображений каждого класса)	450 (90 x 5)	450 (90 x 5)
Объём тестового набора (количество изображений каждого класса)	50 (10 x 5)	50 (10 x 5)
Вероятность зеркального отображения по горизонтали	0,5	0,5
Число эпох обучения	[1, 2, …, 160]	[1, 2, …, 160]
Фиксируемый параметр	доля КРО, %	доля КРО, %

Объединение результатов на основе схемы Шортлиффа

Вероятность наступления хотя бы одного из N совместных независимых событий вычисляется по формуле (1).

Для вычисления общей вероятности события будем использовать схему на основе формулы Шортлиффа [8] и формулы (1) для двух событий:

$$P_K = P_{\text{ОЭ}} + P_{\text{РЛ}}\left(1 - P_{\text{ОЭ}}\right), \quad (2)$$

где P_K – вероятность корректного распознавания объекта (доля корректно распознанных объектов (КРО)); $P_{\text{ОЭ}}$ – вероятность корректного распознавания объекта (доля КРО) для оптического канала, $P_{\text{РЛ}}$ – вероятность корректного распознавания объекта (доля КРО) для радиолокационного канала.

Отличием предложенной схемы от схемы Шортлиффа является то, что в ней используется не мера доверия, а вероятность правильной классификации объекта. Однако основные свойства схемы Шортлиффа в (2) сохраняются:

– она симметрична в том смысле, что порядок классификации в каналах не существенен;

– по мере накопления подкрепляющих свидетельств РК движется к определенности, т.е. стремится к 100 %.

В целях повышения информативности результатов экспериментов, значения доли КРО сглажены методом двустороннего скользящего среднего с шириной окна k = 5:

$$MA_t = \frac{1}{2k+1}\sum_{i=-k}^{k} x_{t+i} = \frac{1}{11}\sum_{i=-5}^{5} x_{t+i}, \quad (3)$$

где MA_t – среднее значение функции в момент времени t; x – значение функции; t – момент времени.

Результаты экспериментов № 1 и № 2 и комплексная оценка доли КРО, вычисленная по формуле (2), после применения к ним формулы (3), представлены на рис. 6: результаты эксперимента № 1 представлены графиком $\mathrm{P_{ОЭ}}$, результаты эксперимента № 2 – графиком $\mathrm{P_{РЛ}}$, комплексная оценка – графиком $\mathrm{P_{\Sigma}}$.

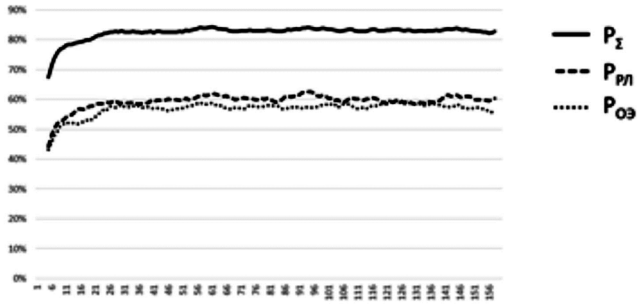


Рисунок 6. Зависимость доли КРО от числа эпох обучения при объединении решений по схеме Шортлиффа

Как видно из графиков, представленных на рисунке 4, при 60 эпохах обучения рост доли КРО прекращается и все три кривых выходят на плато.

Среднее арифметическое значение доли КРО для числа эпох обучения в диапазоне от 60 до 160, полученное в ходе проведения экспериментов, составляет: 58 % для оптического канала (эксперимент № 1), 60 % для радиолокационного канала (эксперимент № 2), 83 % для комплексированной оценки доли КРО. Комплексированная оценка доли КРО на 43 % выше оценки на основе данных только оптического канала и на 38 % выше оценки на основе данных только радиолокационного канала.

Заключение

Эффективность СТЗ наземного БТС, использующей несколько каналов получения информации, может быть значительно повышена путём комплексной обработки данных, получаемых через каждый из каналов. В статье предложен алгоритм принятия решения о принадлежности объекта к определённому классу на основании информации, полученной через оптический и радиолокационный каналы. Предложенный алгоритм с применением схемы Шортлиффа позволяет повысить эффективность СТЗ БТС на 43 % по сравнению с использованием только оптического канала и на 38 % по сравнению с использованием только радиолокационного канала.

Литература

1. Петрова Л. А., Кузнецова Т. Е. Цифровые технологии в экономике и бизнесе. Этап: экономическая теория, анализ, практика. № 2. 2020. с. 74-89. doi: 10.24411/2071-6435-2020-10014

2. И.А. Егоров, Н.С. Шипулин, С.Н. Косников. Современные технологии автоматизации и роботизации процессов аграрного производства. Региональная и отраслевая экономика. 2023 № 2 С. 8-17. DOI: 10.47576/2949-1916_2023_2_8.

3. С.В. Козлов, А.Н. Кубанков, А.П. Шабано. Об информационной поддержке беспилотных воздушных судов на маршрутах железнодорожной логистики. Наукоемкие технологии в космических исследованиях Земли. 2024. Т. 16, №3. С. 4-11. DOI: 10.36724/240954192024163411.

4. Д.С. Чиров, К.В. Новак. Перспективные направления развития робототехнических комплексов специального назначения. Вопросы безопасности. 2018. № 2. С. 50-59. EDN: XRFHFB.

5. Д.С. Чиров, О.Г. Чертова, Т.Н. Потапчук Методика обоснования требований к системе технического зрения робототехнического комплекса. Труды СПИИРАН. 2017. № 2(51). С. 152-176. DOI: 10.15622/sp.51.7.

6. Безумнов Д.Н., Чиров Д.С., Базылев М.В. Алгоритм классификации объектов на основе комплексной обработки данных системы технического зрения наземного робототехнического комплекса. Электросвязь. 2025. № 2. С. 40-45. DOI: 10.34832/ELSV.2025.64.2.006. EDN: RKMFWS.

7. A. Sikdar, S. Cao, Y. F. Zheng, R. L. Ewing. Radar depth association with vision detected vehicles on a highway. 2014 IEEE Radar Conference. DOI: 10.1109/RADAR.2014.6875771.

8. Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project / edited by B.G. Buchanan and E.H. Shortliffe. Addison-Wesley, 1984. 769 p.

9. Tai Fei, Dieter Kraus, Abdelhak M. Zoubir. Contributions to Automatic Target Recognition Systems for Underwater Mine Classification. IEEE Transactions on Geoscience Andremotesensing, vol. 53, no.1, 2015, pp. 505-518.