

Технічні науки

УДК 004

Олейник Александр Александрович

ассистент кафедры ПИ

Харьковский национальный университет радиоэлектроники

Олійник Олександр Олександрович

асистент кафедри ПІ

Харківський національний університет радіоелектроніки

Oliinyk Oleksandr

Assistant of the Department of PI

Kharkiv National University of Radio Electronics

Зыбина Екатерина Викторовна

ассистент кафедры ПИ

Харьковский национальный университет радиоэлектроники

Зибіна Катерина Вікторівна

асистент кафедри ПІ

Харківський національний університет радіоелектроніки

Zybina Kateryna

Assistant of the Department of PI

Kharkiv National University of Radio Electronics

**ОБЗОР НЕКОТОРЫХ МЕТОДОВ РЕШЕНИЯ ЗАДАЧИ
РАСПОЗНАВАНИЯ ОБРАЗОВ В ВИДЕОПОТОКЕ
ОГЛЯД ДЕЯКИХ МЕТОДІВ ВИРІШЕННЯ ЗАДАЧІ
РОЗПІЗНАВАННЯ ОБРАЗІВ У ВІДЕОПОТОЦІ
AN OVERVIEW OF SOME METHODS FOR SOLVING THE PATTERN
RECOGNITION PROBLEM IN A VIDEO STREAM**

Аннотация. Развитие техники в современном мире происходит во все более ускоряющемся темпе, что ведет к усложнению и увеличению потребности в автоматизации. Для решения задач автоматизации требуется возможность оценивать изменения в окружающей среде. Любой биологический организм при оценке изменений окружающей среды во многом опирается на свое зрение. По мнению авторов, естественной является попытка наделить машину возможностью распознавать эти изменения аналогичным образом, то есть через зрение. В данной статье будут рассмотрены идеи некоторых методов (а именно: блочные алгоритмы оценки движения, SLAM), применяемых для решения данной задачи.

Ключевые слова: машинное зрение, блочные алгоритмы оценки движения, SLAM.

Анотація. Розвиток техніки в сучасному світі відбувається у все більше прискорювальному темпі, що веде до ускладнення і збільшення потреби в автоматизації. Для вирішення завдань автоматизації потрібна можливість оцінювати зміни в навколишньому середовищі. Будь-який біологічний організм при оцінці змін навколишнього середовища багато в чому спирається на свій зір. На думку авторів, природною є спроба наділити машину можливістю розпізнавати ці зміни аналогічним чином, тобто через зір. У даній статті будуть розглянуті ідеї деяких методів (а саме: блокові алгоритми оцінки руху, SLAM), що застосовуються для вирішення даного завдання.

Ключові слова: машинний зір, блочні алгоритми оцінки руху, SLAM.

Summary. The development of technology in the modern world is happening at an increasingly accelerating pace. Its development leads to a complication and increase in the need for automation. To solve automation

problems, the ability to evaluate changes in the environment is required. When assessing environmental changes, any biological organism relies heavily on its vision. According to the authors, a natural attempt to empower a machine to recognize these changes in a similar way, that is, through vision. This article will consider the ideas of some of the methods (namely: block motion estimation algorithms, SLAM) used to solve this problem.

Key words: *machine vision, block motion estimation algorithms, SLAM.*

Машинное зрение – класс задач в области искусственного интеллекта, а так же связанные с ним технологии получения и обработки изображений для решения разнообразных задач без участия человека [1].

Для корректной работы систем компьютерного зрения необходимы следующие базовые компоненты: камера, конвертер, который будет заниматься оцифровкой полученного изображения, контроллер управления движением. Опишем процесс работы такой системы. На первом этапе она ожидает получения изображения объекта. На втором этапе происходит оцифровка изображения, его анализ и принятие решений на основе полученных результатов. На последнем этапе передаются указания контроллеру [9].

Машинное зрение – достаточно большое направление в науке и технике. Хотелось бы более подробно рассмотреть такой его подраздел как оценка движения. В данном контексте под информацией о движении понимают двумерный массив векторов, которые характеризуют движение заданных объектов. Существует несколько блочных алгоритмов оценки движения: полный перебор, шаблонные методы, методы иерархического поиска, методы векторов кандидатов и другие [9].

Как понятно из названия, метод полного перебора основан на полном переборе всех возможных блоков изображения и всех возможных векторов их движения. Шаблонные методы предполагают, что кроме

видеопоследовательности имеются еще и шаблоны искомых объектов. Соответственно на каждом шаге мы ищем объект соответствующий шаблону и оцениваем его сдвиг относительно предыдущего состояния. Методы иерархического поиска основаны на следующей идее: два последовательных кадра уменьшаются с определенным шагом до определенного размера. Оценка движения проводится на всех полученных изображениях каким-нибудь известным методом. Процесс начинается с самого маленького кадра и идет к самому большому, на каждом этапе производится уточнение полученного набора векторов. Методы векторов-кандидатов базируются на идее, что вектора движения соседних блоков сонаправлены и, таким образом, в вычислении вектора движения очередного блока изображения участвуют вектора соседних блоков [9].

Рассмотрим и другие методы оценки изменений в окружающей обстановке путем оценки видеопотока. Например, такое семейство методов как SLAM.

Метод одновременной локализации и картирования – метод решения вычислительной проблемы построения или обновления карты неизвестной среды, с одновременным отслеживанием местоположения объекта в ней [2].

Распространено несколько различных реализаций данного метода, имеющих свои особенности: EKF-SLAM, FastSLAM, DP-SLAM.

DP-SLAM (от англ. Distributed Particle - Simultaneous Localization And Mapping) - один из подходов решения задач SLAM, который реализует простой фильтр частиц над картами и положениями робота. Данный метод основан на применении технологии, называемой отображение распределённых частиц. Его удобно применять, когда в исследуемой среде нет возможности найти ярко выраженные ориентиры[3].

EKF SLAM - это усовершенствованный фильтр Калмана (EKF) для синхронной локализации и отображения. Как правило, алгоритмы EKF

SLAM основаны на функциях и используют алгоритм максимального правдоподобия для связи данных [4].

FastSLAM – в основе алгоритма кроме фильтра Калмана и сети Баеса лежит следующая идея: если разбить всю карту на множество подкарт это позволит исключить взаимные связи ключевых точек, а значит, значительно ускорит обработку [5].

Так же хотелось бы хоть немного рассказать о таком классе алгоритмов, как генетические алгоритмы. Они могут, как использоваться для распознавания образов (например, в [6] рассматривается подход автоматического выделения важных блоков на начальном этапе обработки набора изображений) так и участвовать в оценке движения (например, в [7] описывается метод, использующий генетические алгоритмы для распознавания эмоций человека).

Генетический алгоритм - это алгоритм, решающий задачи оптимизации и моделирования с помощью случайного подбора и не менее случайного изменения значений параметров, которые необходимо найти. В его основе лежит идея аналогичная естественному отбору [8, с. 2].

В данной статье были рассмотрены существующие алгоритмы, которые решают разные задачи, возникающие при попытке ответить на вопрос: «Возможно ли научить машину реагировать на зримые изменения окружающей среды?». Как выяснилось данная задача – не так проста, как может показаться на первый взгляд. Существует довольно большой список подзадач, которые возникают, стоит лишь немного углубиться в тему. Например: как нам вообще анализировать видео поток? Как посчитать объекты на изображении? Как нам определить если не что это за объект, то хотя бы к какому классу нам стоит его отнести? Что делать, если изображение загрязнено и не удастся четко определить, к какому классу относится конкретный объект? Что делать если объекты расположены

слишком близко и невозможно изменить угол съемки? Как вообще выделить объект аморфной формы? И т.д.

Современные ученые выбирают себе какой-нибудь вопрос из этих подзадач и пытаются его решить. Среди не решенных можно выделить следующие, перспективные, на мой взгляд, проблемы: «поиск целевого объекта в зашумленном окружении» и «поиск угроз в видеопотоке».

Рассмотрим проблему «поиска целевого объекта в зашумленном окружении» более подробно. Для начала определимся с тем, что такое зашумленное окружение. Это такое состояние видеопотока, когда в нем сложно выделить объекты вообще. Оно может быть вызвано как наличием в кадре каких-нибудь естественных проблем, вроде дыма, песка и т.д., так и просто наличием в кадре огромного количества однотипных объектов, которые могут закрывать искомый объект от обнаружения. Классические методы анализа видеопотока, рассмотренные в начале данного раздела, совершенно не подходят для решения этой задачи. Мало того что они очень сложные и ресурсоемкие, так еще и очень сильно усложняются при работе с изображениями с большим количеством объектов или иных шумов.

Алгоритмы SLAM, которые тоже служат анализу видеопотока, предлагают решать эту задачу методом построения карты окружающего пространства. Основная парадигма – такая: оказавшись в неизвестном пространстве, мы просто будем строить его карту, обследуя его целиком. И, в конце концов, мы найдем интересующий нас объект. Если же пространство нам известно, то мы попробуем уточнить карту, вдруг что-то изменилось, и заодно найдем интересующий нас объект. Этот подход – однозначно хорош, однако не факт, что позволит решить интересующую нас задачу. Он требует дополнительных активных действий платформы, что не всегда возможно.

Методы подсчета численности объектов – хороши для того что бы выделить все объекты, которые присутствуют в кадре, работают достаточно быстро, не требуют лишних движений совершаемых платформой. Однако, они совершенно не позволяют найти один конкретный искомый объект среди всех выделенных.

Очень перспективными для преодоления зашумленности окружающего пространства выглядят эволюционные алгоритмы. Они позволяют анализировать кадр и с большой долей вероятности найти на нём искомый объект, поскольку из-за своей структуры они позволяют искать не только полное соответствие, но и просто что-то похожее на искомый объект. В доказательство можно привести работу, в которой компьютер научили распознавать человеческие эмоции по изображению лица[10]. А они хоть и имеют ряд объединяющих особенностей, практически всегда имеют и отличия в отображении у каждого человека.

Конечно, наиболее перспективным выглядит синтез нескольких алгоритмов: механизма передвижения из SLAM для подводных объектов, механизма подсчета объектов на изображении, для выделения всех объектов, нужного нам класса, и эволюционных алгоритмов, для поиска непосредственно искомого объекта, среди всех обнаруженных.

Литература

1. Мухамедияров Р.М. Машинное зрение: понятия, задачи и области применения. URL: http://www.rusnauka.com/25_NPM_2009/Informatica/50975.doc.htm
2. Durrant-Whyte H., Bailey T. Simultaneous localization and mapping: part I // IEEE Robotics & Automation Magazine. 2006. No 13(2). PP. 99–110. CiteSeerX 10.1.1.135.9810. doi: 10.1109/mra.2006.1638022. ISSN 1070-9932

3. Элиазар А., Парр Р. DP-SLAM: быстрая, надежная одновременная локализация и сопоставление без предопределенных ориентиров.
4. Монтемерло М., Thrun S., Koller D., Вегбрайт Б. FastSLAM: факторизованное решение проблемы одновременной локализации и отображения [pdf] // Труды AAAI Национальной конференции по искусственному интеллекту. 2002. С. 593–598.
5. Алгоритмы локальной навигации и картографии для бортовой системы управления автономного мобильного робота // CyberLeninka. URL: <https://cyberleninka.ru/article/n/algoritmy-lokalnoy-navigatsii-i-kartografii-dlya-bortovoy-sistemy-upravleniya-avtonomnogo-mobilnogo-robota> (дата обращения: 12.02.2021)..
6. Buckles B., Petry F., Prabhu D., Lybanon M. Mesoscale Feature Labeling from Satellite Images / Edit. by Sankar K. Pal, Paul P. Wang // Genetic Algorithms for Pattern Recognition CRC Press, Inc, 1996.
7. Федотов Д.В., Сидоров М.Ю. О применении эволюционных алгоритмов для настройки нейронной сети при решении задачи распознавания эмоций // Актуальные проблемы авиации и космонавтики. 2016. №1.
8. Mitchell Melanie an Introduction to Genetic Algorithms. Cambridge, MA: MIT Press. 1996. [ISBN 9780585030944](https://doi.org/10.1017/CBO9780585030944).
9. nure.ua. URL: <https://nure.ua/wp-content/uploads/workshop/nov-zbirnyk-ist-2018.pdf#4> (дата обращения: 12.02.2021). С. 302-304.
10. О применении эволюционных алгоритмов для настройки нейронной сети при решении задачи распознавания эмоций // cyberleninka.ru. URL: <https://cyberleninka.ru/article/n/o-primenenii-evolyutsionnyh-algoritmov-dlya-nastroyki-neyronnoy-seti-pri-reshenii-zadachi-raspoznavaniya-emotsiy/viewer> (дата обращения: 12.02.2021).