

Соколов К.А.

Національний технічний університет України

«Київський політехнічний інститут імені Ігоря Сікорського»

ВІЗУАЛЬНЕ ВИЯВЛЕННЯ ТА ВІДСТЕЖЕННЯ МАЛОГАБАРИТНИХ РУХОМИХ ОБ'ЄКТІВ НА ОСНОВІ ФУНКЦІОНАЛЬНИХ ОСОБЛИВОСТЕЙ ЗОРОВОГО АПАРАТУ ТА ОСОБЛИВОСТЯХ СПРИЙНЯТТЯ ЛЮДИНИ

У даній статті розглядається питання виявлення і відстеження малогабаритних літальних рухомих об'єктів на основі функціональних особливостей зорового апарату і особливостей сприйняття людини візуальної інформації. Автор розглядає, порівнює різні методи виявлення малорозмірних об'єктів, зокрема малогабаритних безпілотних літальних апаратів (МБПЛА) на відеопослідовності у різних умовах. Зокрема розглянуто принципи виокремлення фігури від фону, тобто виокремлення об'єкту спостереження на сцені з поміж фонової обстановки. Автор наводить паралелі з огляду функціонування та будови людського ока, теорії розпізнавання образів людиною і елементами комп'ютерного зору, що реалізує ці принципи у сфері цифрової обробки зображень та сигналів. Наведено структурну схему будови сучасного відеодавача на прикладі складової частини сучасних систем відеоспостереження – IP камери, як джерела отримання інформації щодо візуальної обстановки. Також автор аналізує можливості використання спеціальних підходів і технологій обробки зображень для поліпшення якості візуального представлення, як етапу передобробки загального методу. Запропоновано кілька евристичних підходів при вирішенні задач попередньої обробки з використанням сегментації та кластеризації зображень. Окремим підрозділом рукопису висвітлено можливість застосування автоматизованих систем керування та описано важливість роботи оператора у системах спостереження, збору та аналізу даних, зокрема візуальної інформації. Наведено основні групи методів детекції руху на відеопослідовностях зображень, проаналізовано їх переваги та недоліки роботи у різних умовах, зазначено особливості використання та застосування. Розглянуто можливість застосування нейронних мереж на етапі постобробки у процесі виявлення МБПЛА задля попередньої класифікації виявленого об'єкту на прикладі роботи сучасного алгоритму YOLOv8. Результати роботи можуть бути використані в наукових дослідженнях, у сферах безпеки та моніторингу та інших областях, де необхідно виявляти і відслідковувати малогабаритні рухомі об'єкти.

Ключові слова: детекція руху, виявлення образів, особливості зору, відеопослідовність, обробка зображень, відслідковування, відеодавач, камера, нейронні мережі, МБПЛА.

Постановка проблеми. В зв'язку з всебічним розвитком та впровадження у різні сфери життя малих безпілотних літальних апаратів питання контролю, детекції, відслідковування та супроводження є надзвичайно важливим. Кількість таких пристроїв невідомо зростає, як і зростає їх сфера застосувань від промисловості і розважальної сфери закінчуючи сферами безпеки.

Здешевлення відеокамер, зменшення їх габаритів дозволяє будувати цілі мобільні системи відстеження та аналізу даних. В менш крупному масштабі системи оператор-дрон достатньо для того щоб проводити безліч операцій та маніпуляцій. З метою забезпечення взаємодії різних систем та забезпечення безпеки повітряного простору, як масштабах підприємств та виробництва, так і в масштабах приватної власності існує необхід-

ність у створенні систем та алгоритмів, що здатні відслідковувати МБПЛА у повітряному просторі та забезпечувати їх супроводження з можливістю подальшої класифікації таких об'єктів. Останній пункт можуть непогано забезпечити штучні нейронні мережі, які перебувають у фазі активного розвитку [1]. А також існує необхідність покращення та оптимізації вже існуючих систем і алгоритмів, які наразі потребують значних витрат різноманітних ресурсів.

Аналіз останніх досліджень і публікацій. Сучасні дослідження дедалі глибше занурюються у деталі функціонування людського ока, периферичного зору зокрема, хоча в цій області вже немало зроблено [2–4]. Поруч активно розвиваються дослідження у сфері сприйняття та обробки візуальної інформації людиною [5–7].

Вищезазначені дослідження також впливають і на розвиток сучасного напрямку комп'ютерного зору, цифрової обробки зображень [8, 9] та штучних нейронних мереж [10]. Важко встановити де проходить конкретна межа між сферами, що призводить до надзвичайно цікавих наукових результатів, виникають роботи, які об'єднують в собі досягнення з різних областей [11, 12] та екстраполюють нові підходи у інші сфери. На базі цих досліджень базується фундаментальна практика використання новітніх «розумних» систем відеоспостереження [13, 14], які здатні виконувати широкий спектр завдань. Зроблено у цих областях немало, новизна полягає у комбінуванні та застосуванні знань з різних областей.

Метою статті є формулювання основних принципів та синтез на їх основі оптимального алгоритму покрокової обробки відеопослідовності задля виявлення та попередньої оцінки малогабаритних літальних об'єктів.

Виклад основного матеріалу. Комп'ютерний зір вивчає питання побудови моделей, аналізу відео та зображень отриманих з відеодавача [15]. Сучасні методи розв'язання задач комп'ютерного зору включають підхід на основі нечіткої логіки; підхід на базі згорткових нейронних мереж та глибокого навчання; підхід з використанням детекторів і дескрипторів [16, 17]. Цей напрямок швидко прогресує, що пов'язано зі збільшенням обчислювальної потужності комп'ютерів, а також залученням великої кількості спеціалістів до досліджень. Комп'ютерний зір використовується в десятках галузей, наприклад, при побудові «розумних» рішень, ідентифікації об'єктів, автоматизації процесів з використанням дронів, веденню відеоспостереження та багато іншого. Комп'ютерний зір почали вивчати у рамках кібернетики з метою відтворення людського бачення та сприйняття.

Цікавою особливістю функціонування зору людини – це здатність периферійного зору досить швидко та чітко визначати наявність руху у певній зоні, а потім можливість зосередити увагу на цій ділянці з об'єктом центральним зором. Дана особливість визначена як фізіологічними особливостями будови зорового апарату, так і психофізичним сприйняттям та подальшою обробкою отриманої зорової інформації певними ділянками мозку людини [18]. Що у свою чергу було обґрунтовано еволюційним розвитком людини, в умовах агресивного навколишнього середовища та високої концентрації небезпек у повсякденному житті. Варто зазначити, що людина здатна фіксувати рух одночасно у кількох зонах. На рисунку 1 показана

будова людського ока [19], з якої частково впливають особливості сприйняття візуальної інформації людиною.

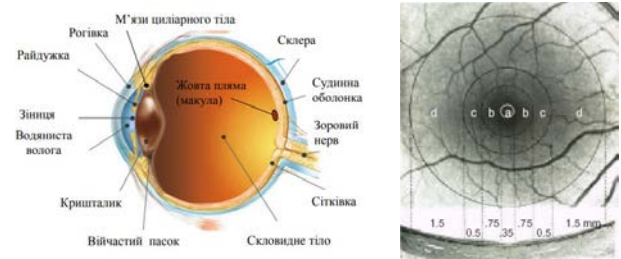


Рис. 1. Будова людського ока та анатомічний вигляд зони жовтої плями: a – foveola, b – fovea, c – parafoveal, d – perifoveal

З цікавих особливостей на рис. 1 можна побачити, що щільність та кількість паличок, по суті, елементарних фотоелементів, що фіксують яскравість у монохромному спектрі, зростають у периферійній ділянці. На ряду з особливостями сприйняття та обробки інформації мозком людини це дає змогу швидко фіксувати рух у ділянках на яких не сфокусована увага, з подальшою можливістю зміщення фокусу на цю ділянку. Детальні дослідження даного факту мають більш глибокі та обширні пояснення [2–4, 20–23].

У відповідність до структури ока та отримання зображення на сітківці а також отримання віртуального зображення внаслідок процесу зорового сприйняття сцени наведемо реалізацію даного механізму у контексті задач комп'ютерного зору [24].

Найпоширенішими та найрозповсюдженішими камерами, які використовуються у побуті, у виробництві, у сферах безпеки є IP камери. Основною відмінністю таких пристроїв від старого покоління аналогових систем відеоспостереження є формат даних на виході таких пристроїв [14]. Такі камери видають на виході цифровий потік даних, який піддається подальшій обробці програмними засобами. Принципова модульна будова такого пристрою показана на рис. 2.

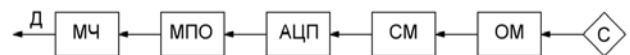


Рис. 2. Структурна схема отримання відеопотоку

На рисунку 2 основні елементи: С – сцена, ОМ – оптичний модуль, СМ – сенсорний модуль, АЦП – аналого-цифровий перетворювач, МПО – модуль процесорної обробки, МЧ – мережева частина, Д – дані.

З огляду на основні елементи сучасної відеокамери спостереження зупинимося на основних параметрах, що відіграють найбільшу роль при

обробці відеопотоку. Узагальнена структурна схема на рис. 2 надана в спрощеному вигляді, але відображає принципові конструктивні ділянки необхідні для отримання потоку відеоданих. З огляду на це упустимо велику кількість конструктивних параметрів і зупинимось на принципово важливих параметрах в рамках даного дослідження.

До вимог відеодавачів (оптичних камер), які можуть використовуватися в системах належать такі параметри як роздільна здатність, час спрацювання затвору, кількість кадрів у секунду, кути огляду (фокусна відстань) та інші. Кожен параметр визначає для яких умов призначена конкретна камера [13].

Роль оператора людини в системах моніторингу, відеоспостереження, стеження і автоматизованих системах має величезне значення для забезпечення ефективної і надійної роботи систем. Основним завданням оператора є аналіз одержуваної інформації та прийняття обґрунтованих рішень на основі наявних даних. У разі виникнення непередбачених ситуацій або надзвичайних подій, оператор здатний оцінити контекст, зібрати інформацію і вирішити, які заходи необхідно провести. При наявності оператора можна використати так звану ePTZ (electronic pan-tilt-zoom) технологію. Звичайна PTZ камера являю собою керовану рухливу платформу, що здатна повертати камеру з моторизованим об'єктивом у вертикальній та горизонтальній площинах. Технологія ePTZ – також дозволяє здійснювати електронне панорамування, нахил та масштабування, але замість механічного руху вбудована функція PTZ імітує панорамування, нахил і масштабування, якщо роздільна здатність камери дозволяє це робити без істотної втрати якості зображення (рис. 3). Наприклад, в камерах з 4K сенсором при використанні ePTZ користувачі можуть створювати кілька унікальних сцен з роздільною здатністю 1920 на 1080 пікселів. Саме тут себе розкриває особливість людського зору гостро фіксувати рух у зоні периферії.

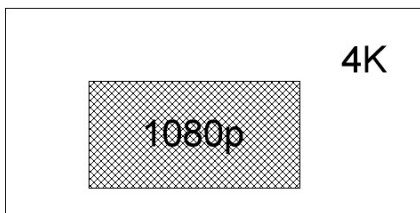


Рис. 3. Вікно функції ePTZ

Збільшення зони огляду цілком логічно підводить до якісного переходу з систем обмеженого кадру до панорамної системи, яка складається

з кількох камер, що здатні забезпечити повний огляд простору на 360 градусів (рис. 4). Такі системи на противагу основній перевазі мають деякі технічні ускладнення у побудові та обробці зображення [8, 9]. Необхідно коректно виконати «зшиття» синхронізованих зображень з різних відеодавачів, компенсувати дисторсії та вирівняти контрастність. Після такої обробки можна отримати придатне для подальшої роботи зображення.



Рис. 4. Формування панорамного зображення

Сегментація зображень також важливий інструмент на шляху до оптимальної обробки зображень. Сучасні камери мають значний вбудований функціонал передобробки зображень. ROI (region of interest) і є прикладом такого функціоналу, що здатний знімати різні частини зображення з різною роздільною здатністю, покращуючи зображення у певній області та знижуючи якість зображення у іншій. Наприклад, при зйомці ділянки вуличної дороги, нижня половина (з автомобілями і дорогою) буде записана в максимальній якості, інша частина зображення (небо, будинки) записується з низькою роздільною здатністю. Або навпаки, небо, як зона спостереження записується з високою роздільною здатністю об'єкти розташовані поблизу землі – низькою. Таким чином, обробка зображення буде здійснюватися на інформативній частині зображення, а бекграунд записується «замиленим» з огляду на низьку інформаційну ємність та для доповнення загальної картини (рис. 5).



Рис. 5. Вікно ROI

Загалом, сучасні системи відеоспостереження вже мають значний широкий функціональний діапазон закладений в програмному забезпеченні кожного пристрою, який вирішує багато типових завдань [13, 14].

Іншою цікавою особливістю сприйняття людиною зорової картини – це інтерпретація та автоматичний поділ сцени на фон та об'єкт. Це було визначено у класичній роботі [25], де надано означення фону та об'єкту (фігури). Дана особливість використана у методах відстеження руху в рамках задач комп'ютерного зору (метод моделювання фону, наприклад).

Існують основні закони сприйняття зображення для виділення окремих об'єктів на зображенні, на яких базується подальше розпізнавання образів [26]. Серед них наступні: принцип фігури-фону, принцип гарної фігури (простоти), принцип близькості, принцип спільності (подібності). Дані принципи фігурують також у інших назвах та доповненнях. Загалом вони пояснюють як людина виділяє об'єкти з поміж іншого (фону) [5, 6].

Яскравий приклад показано на рис. 6 видно однорідну структуру та її перетворення, після яких ми можемо за деякими ознаками говорити про те, що бачимо об'єкти на фоні. Це досягається або за ознакою близькості сусіда, або за кольором, або за принципом доповнення до простої фігури. Таких прикладів можна навести безліч [7]. У контексті виявлення рухомих об'єктів на відеопослідовностях при наявності груп об'єктів, шумах це вкрай важливе питання, оскільки групу пікселів можна об'єднати та ідентифікувати по-різному.

Базовим поняттям гештальт принципів є поняття простоти, яке говорить, що людина організовує сприймане так, щоб воно було регулярним, впорядкованим, симетричним і простим. Надалі цей принцип було обґрунтовано з точки зору теорії інформації [5, 6]. Мозок людини, які ідеальна теплова машина функціонує на максимальну економію нейронних ресурсів людини, і групування сприйраних об'єктів якраз зменшує ентропію (хаос, невизначеність). Отже, мозок має можливість сприймати та обробляти меншу кількість інформації в одиницю часу. Тому фактично гештальт принципи працюють на зменшення кількості інформаційних одиниць для більш ефективної обробки вхідних сигналів. Така здатність сприйняття людини опирається на набутий досвід і є еволюційно обумовленою.

Виходячи з ідеї розбиття зображення на фігуру і фон реалізовані алгоритми відслідковування на базі метод моделювання фону. Метод моделювання фону полягає в створенні статистичної моделі для фонової сцени і подальшому виділенні рухомих об'єктів шляхом визначення відхилень від цієї моделі. До цього класу відносять, наприклад, наступні методи: альфа-змішування, метод

зворотної хвилі, метод власних зображень, суміш розподілів Гауса, Байєсівське моделювання [27].

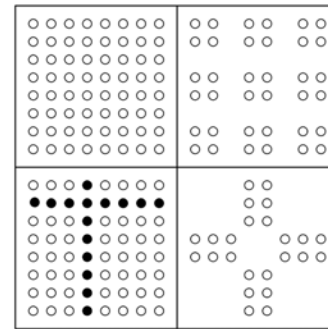


Рис. 6. Сприйняття фігури та фону

Визначення руху за рахунок обчислення диференціальної покадрової різниці – найпростіша реалізація детектору руху на відео, оскільки дозволяє просто та швидко знаходити групи пікселів, які переміщуються з часом на відео.

Формула формування різницевого зображення [27] у загальному вигляді

$$D_i(x, y) = B_i(x, y) - F_i(x, y) \quad (1)$$

де $B_i(x, y)$ – фон

$F_i(x, y)$ – передній план

Алгоритм добре працює в умовах, де фон відносно статичний, та може ефективно виявляти об'єкти, які знаходяться на тлі одного типу (наприклад, небо). Недоліком такого підходу є низька селективність по відношенню до об'єктів руху. Тобто у сценах з динамічними змінами фону, наприклад: рухомими рослинами, коливаннями освітлення і т. д. Тобто, на практиці ми побачимо різноманітні артефакти, рухи дерев, гілок та інших тіл, які не є не цікавими для нас в рамках визначення руху МБПЛА на відеопотоці. Отже основною вимогою є постійне оновлення моделі фону для адаптації до змін.

Приклад роботи такого алгоритму наведено на рисунку 7. Де чорним кольором позначено маску (фон), а білим – об'єкти, що здійснили рух.

Іншою принципово відмінною групою методів детекції руху на відеопослідовності є методи на основі обчислення оптичного потоку. Оптичний потік – це векторне поле, яке представляє напрямки і швидкості руху кожної точки зображення між двома послідовними кадрами. Характерною особливістю цих методів є те, що вони у результаті обробки надають не тільки координати об'єктів руху, а також визначають ще й швидкість переміщення пікселів. Дані методи дозволяють виявляти як глобальні, так і локальні рухи, включаючи швидкі і складні зміни сцени. Також ці методи



Рис. 7. Приклад роботи методу моделювання фону

ефективні при невеликих змінах освітлення. З іншого боку вони можуть бути менш точним у разі великих переміщень або сильних змін освітлення та вимагають наявності текстури або характерних точок на об'єктах для обчислення оптичного потоку. Відеопотік [27] подається у вигляді функції часу та координат пікселів $I(x(t), t)$

$$I(x(t), t) \frac{\partial I(x(t), t)}{\partial t} = \frac{\partial I}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial I}{\partial t} = 0 \quad (2)$$

До методів обчислення оптичного потоку належать: метод Лукаса-Канаде, метод Хорна-Шанка, блокове порівняння та ін. Приклад роботи одного з методів подано на рис. 8.



Рис. 8. Приклад роботи методу визначення руху на основі обчислення оптичного потоку

Покращення та усунення недоліків методів обчислення оптичного потоку є використання дескрипторів особливих точок. Особливі точки – це такі точки які закріплюються на об'єкті для подальшого його відстежування (ознаки). Даних дескрипторів існує значна кількість, які варіюються за принципом обрання цих особливих точок. Це ознаки SIFT, GLOH, MSER та безліч інших [27; 28]. З поміж інших виділяється метод FREAK [29], оскільки він побудований на основі структури сітківки ока.

Також у рамках обробки відеопотоку варто відмітити важливість сегментації та кластериза-

ції зображень (розбиття та об'єднання). Загалом, сегментація зображень дозволяє пришвидшити обробку інформації за рахунок використання алгоритмів паралельного обчислення, тобто віддаючи певні області на обробку певним процесорним потокам. Це дозволяє підвищити швидкість обробки відеопотоку. Також сегментація, яка «по суті» використовується у методі блокового порівняння при обчисленні оптичного потоку, може бути покращена, наприклад, за рахунок введення вагових імовірнісних коефіцієнтів. Дані вагові коефіцієнти обираються емпіричним шляхом виходячи з особливостей конкретної сцени. Наприклад, у випадку МБЛА це може бути очікуваний напрямок підльоту, або зони прольоту. На рисунку схематично показано, що області неба мають більші вагові імовірнісні коефіцієнти, ніж стіна, або асфальтне покриття (рис. 9).

0.98	0.99	0.99	0.99	0.99	0.99	0.99	0.99
0.99	0.95	0.99	0.99	0.99	0.99	0.99	0.99
0.99	0.99	0.99	0.99	0.99	0.99	0.98	0.91
0.99	0.99	0.99	0.89	0.95	0.89	0.89	
0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69
0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69
0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69

Рис. 9. Сегментація за принципом імовірностей появи об'єкта

Іншим класом алгоритмів відслідковування об'єктів на зображенні є методи стеження за об'єктом на основі пошуку за шаблоном. Найпростіший з них – метод шаблонного зіставлення, схематично показаний на рис. 10. Він використовує грубо-точкову стратегію і набір домінуючих ознак для знаходження кандидатів у всьому зображенні. Потім у кожного кандидата ініціалізується локальний процес зсуву середнього значення, і ці

процеси знаходять найближчий локальний максимум у своїх відповідних сусідів.

Ще один метод, який можна використовувати для відстеження об'єкта це кореляційно-екстремальна функція [30]. Кореляційно-екстремальна функція – це метод відстеження об'єктів на відео, який використовує кореляційні фільтри для визначення положення і руху об'єктів в кадрі. Він базується на глобальних характеристиках руху цілі у відео та застосовує фільтр Калмана для коригування траєкторії відстеження рухомої цілі. Цей метод дозволяє підвищити точність і надійність відстеження об'єктів на відео.

В методах такого типу важливим нюансом є наявність якісної шаблонної вибірки. Еталон пошуку піддають процедурам ротації та масштабування, для знаходження більш якісних співпадінь на зображенні. Недоліком таких методів є їхня ненадійність в умовах зашумленості зображення та в умовах складного фону.



Рис. 10. Демонстрація принципу метода шаблонного зіставлення

Використання нейронних мереж для виконання задач. Основним недоліком задіяння штучних нейронних мереж є значна кількість обчислювальних ресурсів необхідних для обробки зображень. Частково ця проблема вирішена шляхом залучення хмарних обчислень до обробки інформації. Залишається також питання швидкості взаємодії, оскільки клієнт-серверні системи мають певну затримку.

Для покращення обробки інформації варто залучати аналіз та обробку інформації нейронною мережею на кінцевому етапі, маючи вже необхідні уточнені вхідні дані, що дасть змогу оптимізувати процес та підвищити швидкодію.

Одним із таких прикладів алгоритмів на основі нейронних мереж є алгоритм YOLO (You Only Look Once). Це алгоритм класифікації та виявлення об'єктів на відео. Він використовує згорткову нейронну мережу для виявлення та класифікації об'єктів у режимі реального часу [10]. YOLO

активно розвивається та поєднує у собі попередньо зазначені у статті алгоритми та методи у різних комбінаціях. Наразі, вже існує та функціонує YOLOv8. На рисунку 11 наведений результат роботи YOLOv8 у контексті вирішення задач відслідковування об'єктів у повітряному просторі.



Рис. 11. Приклад роботи YOLOv8

Пов'язуючи дані від нейронної мережі, яка відіграє роль класифікатора, простих алгоритмів виявлення та відслідковування на виході ми отримаємо вектор що містить наступні параметри: швидкість, поточні координати в кадрі, клас цілі, імовірність приналежності до класу.

$$C_i = [x_i, y_i, T_i^N, P_i^T] \quad (3)$$

де x_i, y_i – координати об'єкта в декартовій системі координат в дискретний момент часу

T_i^N – тип об'єкту з заданого пулу в дискретний момент часу

P_i^T – імовірність приналежності об'єкту до певного типу з заданого пулу.

Висновки. Запропоновано оптимальний алгоритм виявлення об'єкта, його оцінки та попереднього відслідковування (рис. 12). Дана реалізація дозволяє попередньо знизити навантаження на обчислювальну систему, а отже економить обчислювальний ресурс для вирішення задачі пошуку та відслідковування малогабаритного літального апарата у відеопотоці. Отриманий вектор стану надає можливість використовувати отриману інформацію у подальшій обробці та вирішенні більш складних задач, наприклад, як уточнююча інформація у системі комплексної обробки даних з дачив різної фізичної природи. Що, у свою чергу відіграє значну роль у масштабуванні системи для виявлення та відслідковування вже кількох об'єктів на відеопослідовності в складних умовах.

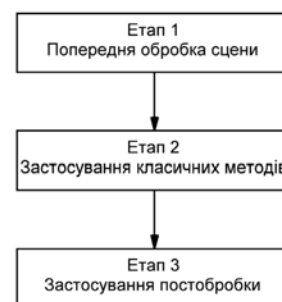


Рис. 12. Блок схема алгоритму виявлення МБПЛА

У запропонованому алгоритмі на першому етапі застосовується попередня обробка сцени виходячи з характерних особливостей самої сцени. На цьому етапі може бути застосований ROI, сегментація з на основі вагових імовірнісних коефіцієнтів та інші методи визначення зон підвищеного інтересу на базі евристичних алгоритмів та функціонування зору людини. Після того, переходимо до наступного етапу – застосування класичних методів. Це групи методів моделювання

фону, обчислення оптичного потоку, методів стеження за об'єктом на основі пошуку за шаблоном та подібні. Останнім етапом постобробки є використання методів, що задіюють нейронні мережі для класифікації попередньо отриманих даних. Така конфігурація дозволяє організувати обробку первинної інформації (відеопотоку) та пришвидшити обробку даних, оскільки кожен попередній етап оптимально обробляє та підготовлює дані для наступного.

Список літератури:

1. Тимошук П. В. Штучні нейронні мережі : навч. посіб. Львів : Вид-во Львів. політехніки, 2011. 444 с.
2. Functional connectivity in the retina at the resolution of photoreceptors / G. D. Field et al. *Nature*. 2010. Vol. 467, no. 7316. P. 673–677. URL: <https://doi.org/10.1038/nature09424> (date of access: 20.09.2023).
3. Poggel D. A., Strasburger H., MacKeben M. Cueing Attention by Relative Motion in the Periphery of the Visual Field. *Perception*. 2007. Vol. 36, no. 7. P. 955–970. URL: <https://doi.org/10.1068/p5752> (date of access: 20.09.2023).
4. Strasburger H., Rentschler I., Juttner M. Peripheral vision and pattern recognition: A review. *Journal of Vision*. 2011. Vol. 11, no. 5. P. 13. URL: <https://doi.org/10.1167/11.5.13> (date of access: 20.09.2023).
5. A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. / J. Wagemans et al. *Psychological Bulletin*. 2012. Vol. 138, no. 6. P. 1172–1217. URL: <https://doi.org/10.1037/a0029333> (date of access: 20.09.2023).
6. A century of Gestalt psychology in visual perception: II. Conceptual and theoretical foundations. / J. Wagemans et al. *Psychological Bulletin*. 2012. Vol. 138, no. 6. P. 1218–1252. URL: <https://doi.org/10.1037/a0029334> (date of access: 20.09.2023).
7. Gregory R. L. *Eye and Brain: The Psychology of Seeing - Fifth Edition*. Princeton University Press, 2015. 296 p.
8. Gonzalez R. C. *Digital image processing*. 3rd ed. Upper Saddle River, N.J : Prentice Hall, 2008. 954 p.
9. Burger W., Burge M. J. *Digital Image Processing*. Cham : Springer International Publishing, 2022. URL: <https://doi.org/10.1007/978-3-031-05744-1> (date of access: 20.09.2023).
10. Real-Time Onboard Object Detection for Augmented Reality: Enhancing Head-Mounted Display with YOLOv8 / M. Łysakowski et al. 2023 IEEE International Conference on Edge Computing and Communications (EDGE), Chicago, IL, USA, 2–8 July 2023. 2023. URL: <https://doi.org/10.1109/edge60047.2023.00059> (date of access: 01.10.2023).
11. Cao F. Application of the Gestalt principles to the detection of good continuations and corners in image level lines. *Computing and Visualization in Science*. 2004. Vol. 7, no. 1. P. 3–13. URL: <https://doi.org/10.1007/s00791-004-0123-6> (date of access: 20.09.2023).
12. Kootstra G., Kragic D. Fast and bottom-up object detection, segmentation, and evaluation using Gestalt principles. 2011 IEEE International Conference on Robotics and Automation (ICRA), Shanghai, China, 9–13 May 2011. 2011. URL: <https://doi.org/10.1109/icra.2011.5980410> (date of access: 20.09.2023).
13. Kolekar M. H. *Intelligent Video Surveillance Systems: An Algorithmic Approach*. Taylor & Francis Group, 2018. 180 p.
14. Nilsson F., Axis C. *Intelligent Network Video: Understanding Modern Video Surveillance Systems, Second Edition*. Taylor & Francis Group, 2016. 366 p.
15. Вовк С. М., Гнатушенко В. В., Бондаренко М. В. *Методи обробки зображень та комп'ютерний зір : навч. посіб.* Дніпро : “ЛІРА”, 2016. 148 с.
16. Сучасні підходи до розв'язання задач комп'ютерного зору / Р. М. Тимчишин та ін. *Control Systems and Computers*. 2018. № 6 (278). С. 46–73. URL: <https://doi.org/10.15407/usim.2018.06.046> (дата звернення: 20.09.2023).
17. Пуятін Є. П., Гороховатський В. О., Матат О. О. *Методи та алгоритми комп'ютерного зору : навч. посіб.* Харків : СМІТ, 2006. 236 с.
18. Іонов І. А., Комісова Т. Є. *Фізіологія сенсорних систем : метод. рек.* 2-ге вид. Харків : ФОП Петров В.В., 2018. 45 с.
19. Сокурєнко В. М., Тимчик Г. С., Чиж І. Г. *Око людини та офтальмологічні прилади [Текст] : навч. посіб.* Київ : НТУУ «КПІ», 2003. 264 с.

20. Mckee S. P., Nakayama K. The detection of motion in the peripheral visual field. *Vision Research*. 1984. Vol. 24, no. 1. P. 25–32. URL: [https://doi.org/10.1016/0042-6989\(84\)90140-8](https://doi.org/10.1016/0042-6989(84)90140-8) (date of access: 20.09.2023).
21. Basler A. Über das Sehen von Bewegungen. *Pflüger, Archiv für die Gesamte Physiologie des Menschen und der Thiere*. 1906. Vol. 115, no. 11-12. P. 582–601. URL: <https://doi.org/10.1007/bf01677292> (date of access: 20.09.2023).
22. Basler A. Über das Sehen von Bewegungen. *Pflüger, Archiv für die Gesamte Physiologie des Menschen und der Thiere*. 1908. Vol. 124, no. 6-8. P. 313–335. URL: <https://doi.org/10.1007/bf01677334> (date of access: 20.09.2023).
23. Solomon J. A., Sperling G. 1st- and 2nd-order motion and texture resolution in central and peripheral vision. *Vision Research*. 1995. Vol. 35, no. 1. P. 59–64. URL: [https://doi.org/10.1016/0042-6989\(94\)e0077-x](https://doi.org/10.1016/0042-6989(94)e0077-x) (date of access: 20.09.2023).
24. Forsyth D. *Computer vision: A modern approach*. Upper Saddle River, N.J : Prentice Hall, 2003. 693 p.
25. Rubin E. *Synsoplevede figurer*. Koebenhavn : Gyldendal, 1915. 228 p.
26. Довбиш А. С., Шелехов І. В. *Основи теорії розпізнавання образів : навч. посіб. у 2-х ч.* Суми : СумДУ, 2015. 109 с.
27. Варфоломєєв А. Ю. *Методи та алгоритми автоматичного відслідковування об'єктів на відеопослідовностях: монографія / ред. О. І. Антонюк.* Київ : КІМ, 2013. 112 с.
28. Leutenegger S., Chli M., Siegwart R. Y. BRISK: Binary Robust invariant scalable keypoints. 2011 IEEE International Conference on Computer Vision (ICCV), Barcelona, Spain, 6–13 November 2011. 2011. URL: <https://doi.org/10.1109/iccv.2011.6126542> (date of access: 20.09.2023).
29. Alahi A., Ortiz R., Vandergheynst P. FREAK: Fast Retina Keypoint. 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Providence, RI, 16–21 June 2012. 2012. URL: <https://doi.org/10.1109/cvpr.2012.6247715> (date of access: 20.09.2023).
30. Overview and methods of correlation filter algorithms in object tracking / S. Liu et al. *Complex & Intelligent Systems*. 2020. URL: <https://doi.org/10.1007/s40747-020-00161-4> (date of access: 20.09.2023).

Sokolov K.A. VISUAL DETECTION AND TRACKING OF SMALL MOVING OBJECTS BASED ON THE FUNCTIONAL FEATURES OF THE VISUAL APPARATUS AND THE PECULIARITIES OF HUMAN PERCEPTION

This article discusses the detection and tracking of small flying moving objects based on the functional features of the visual apparatus and the peculiarities of human perception of visual information. The author considers and compares various methods of detecting small objects, including small unmanned aerial vehicles (UAVs) in video sequences under different conditions. In particular, the principles of figure separation from the background, i.e., the separation of the observed object on the scene from the background environment, are considered. The author draws parallels with the review of the functioning and structure of the human eye, the theory of human image recognition, and elements of computer vision that implement these principles in the field of digital image and signal processing. A structural diagram of a modern video transmitter is presented using the example of a component part of modern video surveillance systems - an IP camera, as a source of information about the visual environment. The author also analyzes the possibilities of using special approaches and image processing technology to improve the quality of visual representation as a preprocessing stage of the general method. Several heuristic approaches are proposed for solving preprocessing tasks using segmentation and image clustering. A separate section of the manuscript highlights the possibility of using automated control systems and describes the importance of operator work in surveillance systems, data collection and analysis, including visual information. The main groups of motion detection methods in video sequences are presented, their advantages and disadvantages in different conditions are analyzed, features of use and application are noted. The possibility of using neural networks at the post-processing stage in the process of detecting UAVs for preliminary classification of the detected object is considered using the example of the modern YOLOv8 algorithm. The results can be used in scientific research, in security and monitoring spheres, and other areas where it is necessary to detect and track small moving objects.

Key words: motion detection, image recognition, features of vision, video sequence, image processing, tracking, video sensor, camera, neural networks, SDUAV.