

УДК 004.042:004.093

ІДЕНТИФІКАЦІЯ ТА ЛОКАЛІЗАЦІЯ ОБ'ЄКТІВ У ВІДЕОПОТОЦІ З ВИКОРИСТАННЯМ КРИПТОГРАФІЧНИХ МЕТОДІВ

Мисюра Ю. О., Ярмілко А. В.

Черкаський національний університет імені Богдана Хмельницького

Abstract. *A method of solving the problem of object localization in a video frame is proposed. Methods of checking the data integrity are reviewed, the primary localization of an object in a video frame and examples of its application are described.*

Однією з важливих задач аналізу відеопотоку є виявлення об'єктів, які з'являються у кадрі або зникають із зони відеоспостереження з плином часу. При цьому об'єкт може бути статичним (нерухомим) або ж динамічним (тобто рухатися з певною швидкістю). Також об'єкт може бути контрастним (його кольорова гама сильно відрізняється від кольору фонових елементів) або неконтрастним на тлі видової сцени.

Перед виконанням локалізації об'єктів у відеокадрі необхідно пересвідчитися в тому, що відеопотік не був модифікований. Для засвідчення автентичності існують різні методи, які відрізняються розміром надлишкових даних і швидкодією. Найбільш розповсюдженим методом захисту інформації в потоках даних є додавання певної надлишкової інформації – хешу. В залежності від потреб користувача, апаратних можливостей і роздільної здатності відеопотоку надлишкова інформація може або забезпечувати відновлення початкових даних у разі їх несанкціонованої модифікації чи пошкодження, або просто сигналізувати про те, що дані було модифіковано.

Забезпечення можливості відновлення початкової інформації полягає у додаванні хешу, який є початковою інформацією, закодованою деяким чином, наприклад множенням блоків інформації на матрицю-ключ [1]. Прикладом хешу, який не забезпечує відновлення початкової інформації, а лише сигналізує про несанкціоновану модифікацію вхідних даних, є хеш MD5 (Message Digest 5), який є шістнадцятковим числом-ідентифікатором блоку даних [2].

Для реалізації алгоритму і демонстрації його роботи було використано мову Python та бібліотеку Scikit-Image, яка є набором інструментів для роботи з зображеннями, а також візуалізації результатів роботи з ними. Перевагою Scikit-Image є те, що функції даної бібліотеки реалізовані для всіх відомих форматів зображень, що дозволяє програмісту зосередитися на виконанні задач на алгоритмічному рівні, а не на рівні даних [3].

Базовий алгоритм виявлення об'єктів у кадрі відеопотоку має таку структуру:

1. Завантажити фрагмент відеопотоку у пам'ять.
2. Розбити фрагмент відеопотоку на окремі кадри. Зазвичай нам не потрібні усі кадри відеопотоку, оскільки аналіз усіх кадрів без

виключення є неефективним з точки зору швидкодії і використання пам'яті. Тому необхідно обрати потрібну частоту кадрів. Вона залежить від динамічних властивостей об'єкту ми хочемо виявити і локалізувати – статичний чи динамічний. Динамічні об'єкти вимагають більшої частоти кадрів для аналізу, причому що більша швидкість об'єкта, то більша частота кадрів необхідна.

3. Порівняти два зображення між собою і виявити об'єкт (або об'єкти).

Оскільки при роботі з відеопотоком часто актуальним є його аналіз у реальному часі, то необхідно обрати метод порівняння зображень з високою швидкістю. У даному дослідженні для порівняння зображень використано метод абсолютної різниці. Він полягає у тому, що при порівнянні двох зображень відкидаються усі попарно однакові пікселі зображення. Ті області, які залишаються, вважаємо об'єктами, які з'явилися в кадрі.

На практиці цей алгоритм дозволяє виявляти розбіжності між двома зображеннями видової сцени з досить високою точністю. Його ефективність виявляється при аналізі як чорно-білих, так і кольорових зображень, при виявленні як простих поодиноких геометричних фігур, так і об'єктів складної структури та при зниженій контрастності.

Експериментальна перевірка методу передбачала програмну оцінку кадрів з відеопотоку з «еталонним», отриманим до потрапляння нових візуальних об'єктів у зону відеоспостереження. Результати роботи криптографічного алгоритму виявлення і локалізації на зображеннях доданих артефактів представлені на рис. 1–3.

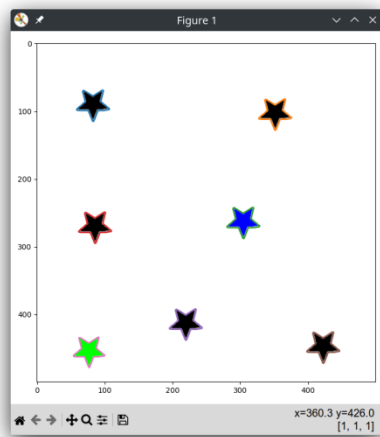


Рисунок 1 – Ідентифіковані та локалізовані об'єкти типу «зірка» з різними спектральними характеристиками на білому тлі

В реальних умовах кадр часто не забезпечує контрастності об'єкту з фоном, ізоморфності як фонових елементів сцени, так і самого об'єкту. Прикладні застосування вимагають виявлення артефактів у природних локаціях (наприклад, на пейзажних сценах), а об'єкт ідентифікації може використовувати засоби маскування. Однак такі об'єкти також добре ідентифікуються методом абсолютної різниці навіть за умови злиття його кольорів і фактур із тлом (рис. 2, 3). Зазначимо, що в наведеному прикладі

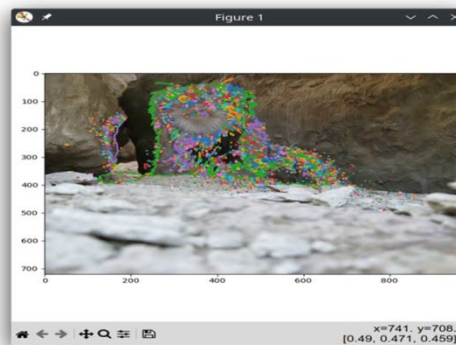
об'єкт (тварина) локалізована кількома окремими графічними кластерами і, фактично, ідентифікація не є повною. Однак факт виявлення зміни попереднього стану видової сцени та зони, в якій такі зміни виявлено, є достатньою інформацією для активації інших спеціалізованих методів обробки відеопотоку та його окремих кадрів.



Рисунок 2 – Пейзаж живої природи, який є тлом видової сцени



а



б

Рисунок 3 – Кадр з об'єктом (твариною), який підлягає локалізації (а) та результат локалізації (б)

Застосування розробленого методу можливе в системах відеоспостереження широкого спектру призначення та в наукових дослідженнях. Зокрема, вбачаються перспективи впровадження у системи пасивного моніторингу квазістатичних сцен та для забезпечення первинних етапів обробки відеопотоку з метою виявлення зон і об'єктів адресного застосування алгоритмів поглибленого інтелектуального аналізу даних.

Література:

1. Yarmilko A., Rozlomie I., Kosenyuk H. (2022) Hash Method for Information Stream's Safety in Dynamic Cooperative Production System. In: Serhiy Shkarlet et al. (Eds): Mathematical Modeling and Simulation of Systems. Lecture Notes in Networks and Systems, vol 344. Springer, Cham. PP. 173-183.
2. RFC 1321 - The MD5 Message-Digest Algorithm : веб-сайт. URL: <https://datatracker.ietf.org/doc/html/rfc1321> (дата звернення: 13.05.2023).
3. scikit-image: Image processing in Python : веб-сайт. URL: <https://scikit-image.org/> (дата звернення: 13.05.2023).

УДК 004.522:004.896

ВИКОРИСТАННЯ VOSK MODEL ДЛЯ ПОКРАЩЕННЯ ЕФЕКТИВНОСТІ РОБОТИ ГОЛОСОВИХ АСИСТЕНТІВ

Науменко С. В., Михайловський П. В.

Черкаський національний університет імені Богдана Хмельницького

Abstract. Using of the Vosk model in order to improve the accuracy of speech recognition, which is critical for the effective work of voice assistants has been discussed in the article. In a particular, the using of the Vosk model in noisy situations, where the accuracy of speech recognition is usually degraded has been looked too.

Голосові асистенти стали популярними у зв'язку з широким поширенням смартфонів та інших пристроїв з голосовим управлінням. Однак, щоб вони могли працювати ефективно, потрібно, щоб точність розпізнавання мовлення була високою.

На сьогоднішній день існує багато голосових асистентів, що пропонують різні функції та можливості. Найбільш відомими серед них є:

1. Google Assistant – голосовий асистент від Google, який працює на різних пристроях з операційною системою Android та iOS. Google Assistant має високу точність розпізнавання мовлення та можливість інтеграції з іншими сервісами Google.

2. Siri – голосовий асистент від Apple, який працює на пристроях з операційною системою iOS та macOS. Siri має великий функціонал та можливість виконання різних завдань, проте має меншу точність розпізнавання мовлення порівняно з Google Assistant [1].

3. Alexa – голосовий асистент від Amazon, який працює на пристроях з операційною системою Amazon Alexa. Alexa має велику кількість функцій та можливостей, проте має деякі обмеження у виконанні завдань поза екосистемою Amazon [2].

Незважаючи на переваги, кожен з цих голосових асистентів має свої недоліки, зокрема: низька точність розпізнавання мовлення у шумних умовах; обмежена кількість мов, на яких можна використовувати асистента; неможливість розуміти специфічний акцент користувача; не завжди точна інтерпретація запиту користувача та надання відповіді на запит; низька можливість інтеграції з певними програмами або сервісами.

Використання бібліотеки Vosk model дозволяє вирішити деякі проблеми, пов'язані з точністю розпізнавання мовлення у шумних умовах та обмежену кількість мов у голосових асистентах. Бібліотека Vosk model використовує моделі глибокого навчання для розпізнавання мовлення, що дає більш точні результати у порівнянні з іншими методами.