

АДАПТИВНИЙ АЛГОРИТМ ВИЯВЛЕННЯ РУХОМИХ ОБ'ЄКТІВ ПІД ЧАС ВІДЕОМОНІТОРИНГУ

© Гривачевський А. П., Прудіус І. Н., 2016

На основі аналізу наявних методів виявлення рухомих об'єктів запропоновано адаптивний до дестабілізуючих факторів алгоритм виявлення, що ґрунтується на паралельному використанні двох методів: методу віднімання фону та методу міжкадрової різниці, а також представлено результати його моделювання. Статистична оцінка параметрів, що виконується в процесі навчання алгоритму, дозволяє використовувати цей алгоритм без налаштування його параметрів практично з будь-якою відеокамерою незалежно від рівня її власних шумів. У результаті роботи запропонованого алгоритму досягається ефективне виділення рухомих об'єктів у їх природних кольорах без малоінформативного фону. Використання цього алгоритму, наприклад, в системах відеоспостереження, дасть змогу спростити сприйняття інформації та зменшити психологічне навантаження оператора.

Ключові слова: відеомоніторинг, виявлення рухомих об'єктів, віднімання фону, міжкадрова різниця, морфологічна фільтрація.

A. Hryvachevskyi, I. Prudyus
Lviv Polytechnic National University

ADAPTIVE ALGORITHM OF MOVING OBJECTS DETECTION IN VIDEO MONITORING

© Hryvachevskyi A., Prudyus I., 2016

The paper based on the analysis of existing methods of detecting moving objects offer combined detection algorithm that adapts to such destabilizing factors as: the noise of the equipment (photoelectron matrix converter, amplifier, analog-to-digital converter), change lighting scenes, change the background. The essence of the proposed algorithm is used two parallel detection methods: the method of background subtraction and frame difference method, which allows the algorithm to work effectively in cases of dynamic background. Statistical evaluation of parameters of the model background, which is performed in the process of learning algorithm for each pixel of the video frame in order to determine the optimal threshold binarization algorithm enables this without setting its parameters with just about any camera, regardless of the level of intrinsic noise eye, which each camera is different. To increase the efficiency of detection of moving objects in the final stage of the algorithm is filtered binary mask by applying morphological operations to remove wrongly marked pixels and combining pixel marked correctly in the objects. The results of the simulation of information processing in the system of video monitoring of mobile objects on a real video signal. As a result of the proposed algorithm is achieved by efficient allocation of moving objects in their natural colors without uninformative background. Using this algorithm, for example, in video surveillance systems, will significantly reduce the amount of data that will be stored in the video archive and will simplify the perception of information by the operator and reduce its psychological burden. The results of the algorithm in the future can be used for high-level analysis of mobile objects, in particular, identifying objects and determining their trajectories.

Key words: video monitoring, moving object detection, background subtraction, frame difference, morphological filtering.

Вступ

Сьогодні дуже поширені системи моніторингу, що використовують як давач сигналу цифрову відеокамеру. До них належать різноманітні диспетчерські системи, системи візуального контролю, а також системи спостереження за об'єктами, розташованими у недоступних місцях [1].

У сучасних прикладних системах відеомоніторингу, в яких використовуються алгоритми відеоаналізу та розпізнавання об'єктів у відео, вкрай важливе правильне виділення рухомих об'єктів [2]. Під час роботи таких систем у реальних умовах (за наявності шумів, тіней, зміни освітлення і заднього плану, під час інтенсивного руху) існуючі методи виділення рухомих об'єктів враховують наявність шуму [3], але не враховують зміну рівня шуму в просторі й часі [4]; вони адаптуються до різкої зміни освітленості або заднього плану, але після цього тривалий час працюють з недостатньою імовірністю виявлення рухомих об'єктів і, як наслідок, некоректними результатами відеоаналізу [3]. Відповідно, необхідно вирішувати завдання дослідження відомих методів виявлення рухомих об'єктів для їх вдосконалення та розроблення нових методів.

Завдання виділення рухомих об'єктів вирішується за два етапи: перший – це виділення пікселів переднього плану, а другий – об'єднання виділених пікселів у об'єкти. В цій роботі пропонується один з варіантів використання такого підходу.

Алгоритм виявлення. Відеосигнал, що формується цифровою відеокамерою, являє собою послідовність відеокадрів $I_1, I_2, \dots, I_K$, кожен з яких описується прямокутною матрицею (растром), яка складається з M рядків та N стовпців. Елементи $I_t[m, n]$ відповідного відеокадру називаються пікселами і представляються як скалярні величини у випадку монохромного зображення.

Здебільшого за відсутності руху на спостережуваній сцені відеокамера системи відеомоніторингу фіксує нерухому сцену, яку називають фоном $BG[m, n]$. Якщо знехтувати впливом на сформоване зображення відеокадру шумів активних елементів системи формування зображення (шуми матриці фотоелектронного перетворювача, підсилювачів, аналого-цифрового перетворювача тощо), а також вважати, що за час формування певної послідовності відеокадрів освітлення та інші умови знімання є незмінними, то за відсутності руху в зоні спостереження, яка відповідає пікселу з координатами $[m, n]$, можна вважати його значення постійним, тобто

$$I_t[m, n] = I_{t-1}[m, n] \text{ для } t = 1, 2, K$$

Отже, постійне значення яскравості піксела $I[m, n]$ у межах певної послідовності відеокадрів розглядається як критерій відсутності руху в відповідній зоні спостережуваної сцени.

На практиці під час формування зображень відеокадрів наявність випадкових завад (шумової складової) неминуча. Основним джерелом таких завад є шум апаратури, похибка квантування у разі перетворення аналогового сигналу в цифрову форму, а також похибки корекції зображення (наприклад, похибка заокруглення під час виконання математичних операцій).

У загальному випадку відеосигнал $I_t[m, n]$ можна розглядати як адитивну суміш істинного значення сигналу $I_t^*[m, n]$ та деякого випадкового шуму $e_t[m, n]$:

$$I_t[m, n] = I_t^*[m, n] + e_t[m, n], \quad (1)$$

де $e_t[m, n]$ – реалізація випадкового шумового процесу в пікселі t -кадру з координатами $[m, n]$.

У випадку, коли відеокамера спостерігає нерухому сцену, істинне значення відеосигналу відповідає фону $BG[m, n]$:

$$I_t[m, n] = BG[m, n] + e_t[m, n]. \quad (2)$$

Враховуючи наявність множини незалежних випадкових факторів, які визначають флуктуації вихідного сигналу відеокамери, можна з достатньою точністю розглядати значення шуму e як випадкової величини, що описується нормальним законом розподілу і характеризується функцією густини розподілу імовірностей:

$$f(e; m, s) = \frac{1}{s \sqrt{2\pi}} \exp\left(-\frac{(e - m)^2}{2s^2}\right) \quad (3)$$

з математичним очікуванням $m = 0$ та із середнім квадратичним відхиленням $s = \text{const}$.

Рух у відеосигналі проявляється в різкій зміні яскравості окремих пікселів, за координатами яких відбувається переміщення об'єкта. Інші коливання яскравості здійснюються плавно, наприклад, зміна природного освітлення упродовж дня.

Припустимо, що в період настроювання системи моніторингу руху, який триває перші K кадрів, відеокамера спостерігає нерухому сцену, яка містить лише статичні об'єкти фону. Це дає змогу достатньо точно виконати оцінку істинного значення яскравості фону:

$$m[m, n] = \frac{1}{K} \sum_{t=1}^K I_t[m, n] \cong BG[m, n] \quad (4)$$

та величину дисперсії шуму $s^2[m, n]$ відеокамери:

$$s^2[m, n] = \frac{1}{K-1} \sum_{t=1}^K (I_t[m, n] - m[m, n])^2. \quad (5)$$

Після завершення оцінки параметрів сцени система переходить у режим пошуку рухомих об'єктів. Як критерій виявлення рухомих об'єктів у межах пікселя з координатами $[m, n]$ можна використати таку умову:

$$|m[m, n] - I_t[m, n]| \geq T, \quad (6)$$

де T – деяке порогове значення.

На вибір величини T впливають різні фактори (величина шуму, контраст рухомого об'єкта щодо фону тощо). Відповідно до властивості нормального закону розподілу, яким описуються шуми камери нашої системи моніторингу, можна вважати, що всі значення пікселів фону з урахуванням шумів повинні потрапляти в інтервал від $T_1 = I_t[m, n] + 3s[m, n]$ до $T_2 = I_t[m, n] - 3s[m, n]$. З урахуванням цієї властивості критерій виявлення руху (6) можна записати так:

$$|m[m, n] - I_t[m, n]| > 3s[m, n]. \quad (7)$$

де $s[m, n]$ – середнє квадратичне відхилення, що визначене в процесі настроювання системи моніторингу руху для кожного пікселя з координатами $[m, n]$.

Для того, щоб алгоритм був адаптивний до зміни фону впродовж часу спостереження для кожного наступного кадру та для кожного пікселя цього кадру обчислюється поточна оцінка середньої яскравості (істинного значення) пікселя [5]:

$$m_{t+1}[m, n] = a m_t[m, n] + (1 - a) I_t[m, n] \quad (8)$$

та величина дисперсії шуму [5]:

$$s_{t+1}^2[m, n] = a \left[s_t^2[m, n] + (m_{t+1}[m, n] - m_t[m, n])^2 \right] + (1 - a) \cdot (I_{t+1}[m, n] - m_{t+1}[m, n])^2 \quad (9)$$

де a – коефіцієнт швидкості навчання (оновлення фону).

Важливо відзначити, що оновлення параметрів фону потрібно здійснювати лише в тих областях, де рух у цей момент часу відсутній. Така особливість алгоритму зумовлена тим, що за високої частоти кадрів відеокамери та повільного руху об'єкта по сцені цей об'єкт поступово за допомогою оновлення буде перенесений у фон.

Для адаптації алгоритму виявлення рухомих об'єктів до динамічного фону паралельно з описаним вище основним методом використовується метод міжкадрової різниці.

Моделювання алгоритму. З метою експериментальної перевірки результату роботи запропонованого алгоритму виконано програмну реалізацію обробки інформації в системі відеомоніторингу рухомих об'єктів у середовищі MATLAB. Тестування алгоритму проводилося на реальному відеопотоці, що отримала камера системи зовнішнього відеоспостереження, з розміром кадру 1280×720 пікселів.

На етапі настроювання системи моніторингу руху, який триває перші $K = 200$ кадрів, алгоритм здійснює оцінку параметрів фону: середнього значення яскравості пікселів та дисперсії шуму (рис. 1).

Після завершення оцінки параметрів сцени система переходить у режим пошуку рухомих об'єктів. У випадку, якщо значення яскравості окремого пікселя $I_t[m, n]$ в поточному відеокадрі виходить за межі $T_1 = I_t[m, n] + 3s[m, n]$ або $T_2 = I_t[m, n] - 3s[m, n]$, то алгоритм відзначає цей піксел, як такий, що належить рухомому об'єкту.

На рис. 1 наведено значення яскравості пікселя з координатами $[120; 210]$ в растрі відеокадру упродовж 2000 кадрів.

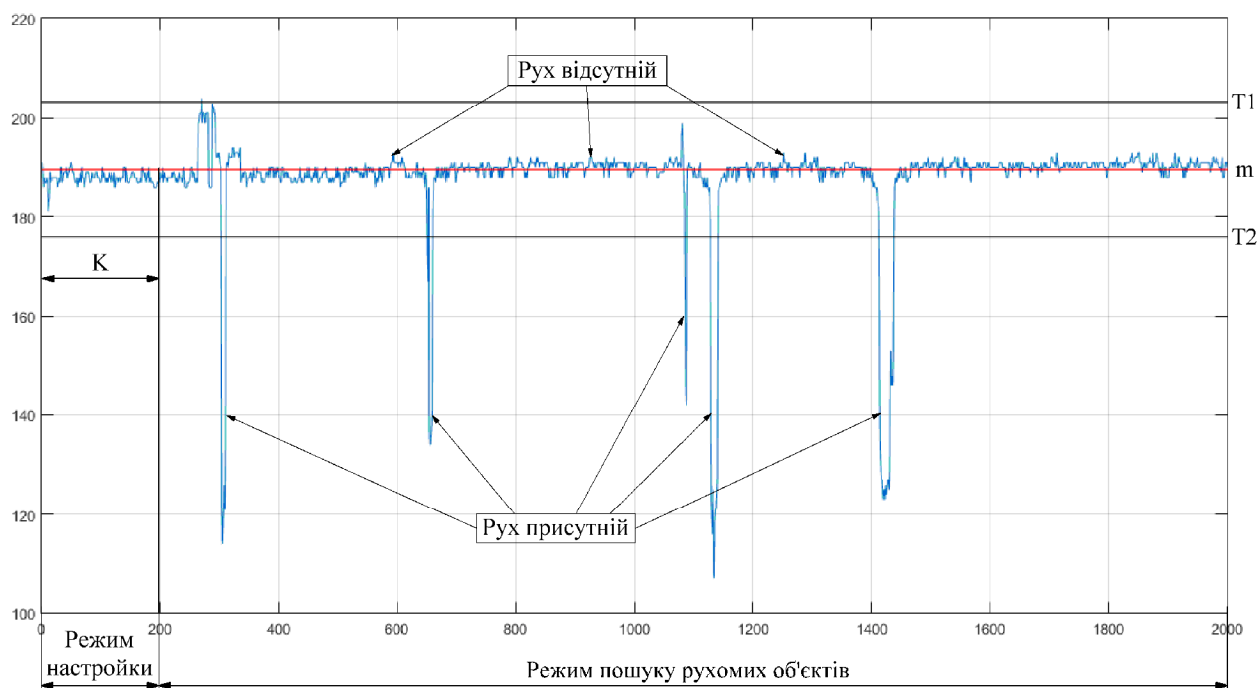


Рис. 1. Значення яскравості пікселя з координатами $[120; 210]$ упродовж 2000 кадрів

У процесі порівняння всіх пікселів поточного відеокадру з їх порогоми здійснюється формування бінарного зображення, в якому “1” відповідає наявності руху в зоні сцени, за яку відповідає цей піксел, а “0” – відсутності. Таке бінарне зображення називається маскою руху. Після формування бінарної маски відбувається її фільтрація з метою видалення помилково відмічених пікселів та об'єднання правильно виділених пікселів у об'єкти. Така фільтрація реалізується за допомогою морфологічних операцій [6].

Сформована результуюча маска надалі використовується для виділення рухомих об'єктів рамкою на поточному кадрі для кращого візуального сприйняття (рис. 2). За необхідності результуюча маска дає змогу виділяти рухомі об'єкти на поточному кадрі в їх природних кольорах та відображати тільки їх без малоінформативного фону (рис. 3).

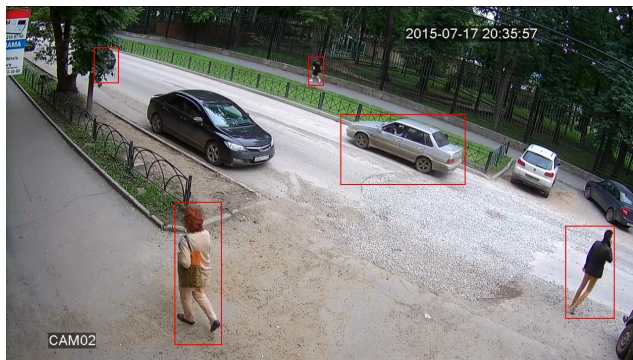


Рис. 2. Кадр №1388 з тестового відеопотоку

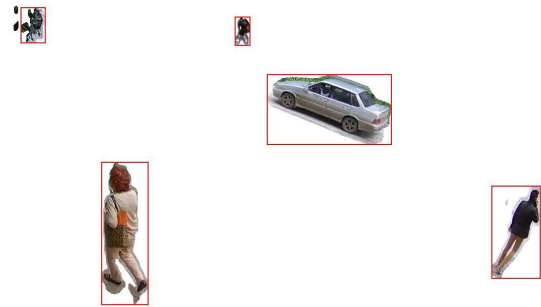


Рис. 3. Виділені рухомі об'єкти в природних кольорах

Висновки

Запропонований адаптивний алгоритм виявлення рухомих об'єктів у режимі реального часу дає змогу ефективно виявляти рухомі об'єкти за наявності таких дестабілізуючих факторів, як: шуми апаратури, зміна освітлення сцени, зміна фону. Статистична оцінка порога бінаризації, яка здійснюється під час роботи алгоритму, дозволяє використовувати цей алгоритм практично з будь-якою відеокамерою, незалежно від рівня її власних шумів, оператор системи відеомоніторингу не повинен емпірично налаштовувати параметри алгоритму, що зменшує вимоги до його кваліфікації. Також спосіб подання результатів виявлення виділенням площі рухомих об'єктів у їх природних кольорах без малоінформативного фону дає змогу значно зменшити об'єм даних, які надходять з відеокамери, що, своєю чергою, дозволяє підвищити ефективність роботи оператора системи відеомоніторингу. Результати роботи алгоритму надалі можна використати для високорівневого аналізу рухомих об'єктів, зокрема, ідентифікації, визначення траєкторії руху об'єктів тощо.

1. Обухова Н. А. Обнаружение и сопровождение движущихся объектов методом сопоставления блоков // Информационно-управляющие системы. 2004. № 1. С. 30–37. 2. Neeti A. A survey of Techniques for Human Detection from Video // University of Maryland. Technical report. 2005. 3. Piccardi M. Background Subtraction Techniques: // A Review IEEE SMC/ICSMC. Vol. 4. 2004. P. 3099–3104. 4. Харебов П., Новиков С. Проблемы выделения объектов в компрессированном потоке изображений // Proceedings of Graphicon. 2009. 5. http://courses.graphicon.ru/files/courses/vision/2012/lectures/cv2012_11_basicvideo.pdf. 6. Шестов Р. В., Тамьяров А. В. Краткая характеристика современных методов морфологической обработки полутоновых изображений в среде Matlab // Вестник ВУиТ. 2013. № 2. С. 91–98.

References

1. N. A. Obukhov, Detection and tracking of moving objects by comparing blocks // Information and Control Systems. 2004. No. 1. P. 30–37. 2. A. Neeti, A survey of Techniques for Human Detection from Video // University of Maryland, Technical report. – 2005. 3. M. Piccardi, Background Subtraction Techniques: A Review IEEE SMC / ICSMC, vol. 4, pp. 3099-3104, 2004. 4. P. Kharebov, Novikov problems selecting objects in a compressed image stream. Proceedings of Graphicon, 2009. 5. http://courses.graphicon.ru/files/courses/vision/2012/lectures/cv2012_11_basicvideo.pdf. 6. R. V. Shestov, A. V. Tamyarov, Brief description of modern methods of morphological processing halftone images among Matlab // Herald VUiT. – 2013. – No. 2. P. 91–98.