

ПОРІВНЯЛЬНИЙ АНАЛІЗ МЕТОДІВ ВЕДЕННЯ ОБ'ЄКТА НА ВІДЕО З КАМЕРИ БЕЗПІЛОТНОГО ЛІТАЛЬНОГО АПАРАТУ

Національний авіаційний університет

artem.chyrkov@nau.edu.ua

Наведено огляд таких методів ведення об'єкта на відео: алгоритм Лукаса – Канаде, алгоритм TLD та підхід, заснований на аналізі локальних гістограм. Виконано порівняльний аналіз якості їх роботи

Ключові слова: ведення об'єкта; алгоритм Лукаса–Канаде; алгоритм TLD; локальні гістограми; порівняльний аналіз

Постановка проблеми

Останнім часом має місце значний прогрес у сфері розробки та практичного застосування безпілотних літальних апаратів (БПЛА). Наприклад, їх можна використовувати для моніторингу ситуації на землі, зокрема для ведення об'єктів, що представляють деякий практичний інтерес (автомобілі або пішоходи для задач патрульної поліції, військова техніка для задач збройних сил тощо). При цьому процес ведення може бути як неавтоматичним (за об'єктом слідкує оператор), так і автоматичним (пошук положення об'єкта на кожному з кадрів відео з камери цільового призначення БПЛА за допомогою деякого алгоритму). Перший спосіб є більш універсальним, але менш надійним через «людський фактор». Другий спосіб є вирішенням даної проблеми, але у свою чергу актуалізує проблеми якості та швидкості: чим якісніше відбувається ведення об'єкта, тим повільніше працює алгоритм, і навпаки.

Задача ведення об'єкта не є специфічною тільки для БПЛА і може виникати в інших областях. Тому на даний момент розроблено декілька підходів до її розв'язання; їх огляд наведено в [1]. Але кожен із методів має свою сферу застосування. Це пов'язано з особливостями кадрів відео, що у свою чергу залежить від конкретної сфери, де виникає дана задача.

Постановка задачі ведення об'єкта така: дано послідовність кадрів з камери БПЛА та виділену область на першому

кадрі, в якій знаходиться об'єкт. Необхідно знайти положення об'єкта на кожному з наступних кадрів.

Основними особливостями, що ускладнюють автоматичну обробку кадрів відео з камери БПЛА (у тому числі ведення об'єкта), є, по-перше, відсутність нерухомого фону (оскільки камера рухається разом із БПЛА, змінюється вся область кадру), і, по-друге, наявність у кадрі (у більшості практичних випадків) текстурованих областей (трава, пісок, ліс, вода тощо), які у більшості випадків мають такий самий вплив на результат роботи алгоритмів, як і шум.

Враховуючи ці особливості, до алгоритму ведення об'єкта ставляться вимоги робастності до шуму і не константного фону. Найбільш часто використовувані підходи, які задовольняють даним вимогам, – клас методів «оптичний потік» (розглянемо на прикладі однієї з реалізацій: алгоритмі Лукаса–Канаде); алгоритм TLD (розглянемо на прикладі його реалізації в бібліотеці OpenCV); методи, засновані на статистичному аналізі (розглянемо на прикладі метода, заснованого на аналізі локальних гістограм, далі – «гістограмний метод»).

Аналіз публікацій

Клас методів «оптичний потік» [2]. Оптичний потік – це представлення видимого сліду руху об'єктів, поверхонь і граней візуальної сцени, що спостерігається під час відносного руху між спостерігачем (наприклад, око людини або ка-

мера) і сцени. Методи розрахунку оптичного руху намагаються оцінити рух між двома кадрами зображення, які відносяться до моментів часу t і $t+\Delta t$ в кожній позиції вокселя (впорядкована трійка значень (x, y, t) , де x і y – координати пікселя, що розглядається; t – час, у випадку послідовності кадрів відео в його якості може використовуватись порядковий номер кадру). Такі методи називаються диференційними, оскільки вони ґрунтуються на локальних оцінках ряду Тейлора для сигналу зображення; тобто вони використовують часткові похідні по просторовим і часовим координатам. Для двовимірного випадку для вокселя (x, y, t) з інтенсивністю $I(x, y, t)$, який буде рухатись на Δx , Δy і Δt між двома кадрами зображення, і з наступним обмеженням сталості яскравості можна записати:

$$I(x, y, t) = I(x + \Delta x, y + \Delta y, t + \Delta t)$$

Вважаючи, що рух був незначним, розраховуючи наближення зображення в $I(x, y, t)$ за допомогою рядів Тейлора можна отримати:

$$I(x + \Delta x, y + \Delta y, t + \Delta t) = I(x, y, t) + \frac{\partial I}{\partial x} \Delta x + \frac{\partial I}{\partial y} \Delta y + \frac{\partial I}{\partial t} \Delta t + \dots$$

З цих рівнянь випливає:

$$\frac{\partial I}{\partial x} \Delta x + \frac{\partial I}{\partial y} \Delta y + \frac{\partial I}{\partial t} \Delta t = 0;$$

в результаті чого:

$$\frac{\partial I}{\partial x} V_x + \frac{\partial I}{\partial y} V_y + \frac{\partial I}{\partial t} = 0;$$

де V_x, V_y – це компоненти x і y швидкості оптичного потоку $I(x, y, t)$, $\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}, \frac{\partial I}{\partial t}$ є похідними для зображення в (x, y, t) у відповідних напрямках.

Таким чином, отримується основне рівняння оптичного потоку:

$$I_x V_x + I_y V_y = -I_t, \text{ або } \nabla I^T \cdot \dot{\mathbf{V}} = -I_t$$

Дане рівняння з двома невідомими не має єдиного розв'язку. Це відомо як проблема апертури алгоритму оптичного потоку. Для розрахунку оптичного потоку необхідно скласти набір додаткових рівнянь, що задаються деякими додатковими

умовами. Всі методи розрахунку оптичного потоку вводять додаткові умови для оцінки фактичного потоку.

Алгоритм Лукаса–Канаде вирішує вищенаведену проблему за рахунок використання інформації про сусідні пікселі в кожній точці [3, 4]. Метод ґрунтується на припущенні, що в локальному околі кожного пікселя значення оптичного потоку однакове, таким чином можна записати рівняння оптичного потоку для всіх пікселів в околі і розв'язати систему рівнянь методом найменших квадратів.

Припустимо, що зміщення пікселів між двома кадрами — відносно мале. Розглянемо піксель p , тоді відповідно до алгоритму Лукаса — Канаде, оптичний потік має бути однаковий для всіх пікселів, які знаходяться у вікні з центром в точці p . А саме, вектор оптичного потоку (V_x, V_y) в точці p має бути розв'язком системи рівнянь:

$$\begin{cases} I_x(q_1)V_x + I_y(q_1)V_y = -I_t(q_1), \\ I_x(q_2)V_x + I_y(q_2)V_y = -I_t(q_2), \\ \dots \\ I_x(q_n)V_x + I_y(q_n)V_y = -I_t(q_n); \end{cases}$$

де q_1, q_2, \dots, q_n – пікселі всередині вікна, $I_x(q_i), I_y(q_i), I_t(q_i)$ – часткові похідні зображення I по координатам x, y і часу t , обчислені в точці q_i .

Розв'язавши систему методом найменших квадратів, отримуємо:

$$\begin{bmatrix} V_x \\ V_y \end{bmatrix} = \begin{bmatrix} \sum_i I_x(q_i)^2 & \sum_i I_x(q_i)I_y(q_i) \\ \sum_i I_x(q_i)I_y(q_i) & \sum_i I_y(q_i)^2 \end{bmatrix}^{-1} \times \begin{bmatrix} -\sum_i I_x(q_i)I_t(q_i) \\ \sum_i I_y(q_i)I_t(q_i) \end{bmatrix}$$

Алгоритм TLD (tracking, learning, detection) [5]. Суть даного алгоритму – об'єднання результатів алгоритмів ведення об'єкта (tracking), знаходження об'єкта (detection) та онлайн-навчання (навчання на кожному отриманому кадрі) (learning). Основна ідея – створити алгоритм знаходження об'єкта, який оновлюється та по-

кращує якість розпізнавання на кожному кадрі.

В якості алгоритму ведення на практиці як правило застосовується алгоритм Лукаса–Канаде. В якості алгоритму знаходження об'єкта – послідовний рандомізований ліс, в якості ознак для нього – двобітні бінарні шаблони.

Для формування вибірки, на якій відбувається онлайн-навчання, вводяться поняття розширювальних та звужуючих подій: розширювальна подія – подія, при настанні якої необхідно внести відповідну область кадру до навчальної вибірки; звужуюча подія – подія, при настанні якої відповідна область кадру до навчальної вибірки не заноситься.

Послідовність обробки поточного кадру така: знайти положення об'єкта за допомогою алгоритму ведення, знайти положення об'єкта за допомогою алгоритму знаходження, об'єднати їх результати, перевірити окіл отриманої області на наявність інших областей, які потенційно можна використати для онлайн-навчання (за допомогою розширювальних та звужуючих подій).

Гістограмний метод [6]. Локальними ознаками для даного метода є гістограма яркостей: гістограма розподілу значень пікселів усередині локальної області. На першому кадрі розраховується гістограма виділеної області, на наступних новим положенням об'єкта вважається область, що має гістограму, найближчу до першої.

На кожному поточному кадрі пошук нового положення об'єкта відбувається шляхом перевірки околу положення на попередньому кадрі. Відстань між гістограмами розраховується з використанням стандартних функціоналів; у даній роботі використовується такий функціонал:

$$dist = \sum_{i=1}^n (h_i^{(1)} - h_i^{(2)})^2,$$

де n – кількість стовпців гістограми, $h_i^{(1)}$ і $h_i^{(2)}$ – гістограми, відстань між якими шукається.

Постановка задачі

Вище було відмічено, що кожен з алгоритмів має свою сферу застосування, в залежності від особливостей відео, характерних для кожної конкретної предметної області. Метою даної статті є виконання порівняльного аналізу алгоритмів Лукаса – Канаде, TLD та гістограмного підходу на відео з різними особливостями.

Викладення основного матеріалу

Для аналізу якості алгоритмів, що розглядаються, використовується розмічена база даних: набір відеофайлів, для кожного з яких на кожному кадрі вказано положення об'єкта, який потрібно вести. Характеристики тестових відео такі:

- відео №1: 1920x1080, 30 кадрів/сек, низький рівень шуму; площа розміченої області 2205 пікселів;
- відео №2: 540x360, 30 кадрів/сек, значний рівень шуму; площа розміченої області 510 пікселів.

Окремо проводиться аналіз «довгострокового» та «короткострокового» ведення. Метод аналізу «довгострокового» ведення: на першому кадрі задається область (в її якості береться розмітка першого кадру); на наступних за допомогою кожного з алгоритмів отримується нове положення області, яке порівнюється з розміткою. Якісною характеристикою алгоритму є близькість отриманих положень області до розмітки, кількісно – середнє арифметичне \bar{d} та середньоквадратичне відхилення S_d відстаней між центром отриманої області та центром розмітки по всім кадрам:

$$d_i = \sqrt{(x_i^{(L)} - x_i^{(T)})^2 + (y_i^{(L)} - y_i^{(T)})^2}, i = 1 \dots N$$

$$\bar{d} = \frac{1}{N} \sum_{i=1}^N d_i; S_d = \frac{1}{N-1} \sqrt{\sum_{i=1}^N (d_i - \bar{d})^2},$$

де $(x_i^{(L)}, y_i^{(L)})$ і $(x_i^{(T)}, y_i^{(T)})$ – відповідно центри розміченої та отриманої областей на i -му кадрі; N – кількість розмічених кадрів.

Метод аналізу «короткострокового» ведення аналогічний вищенаведеному з тою різницею, що область задається заново на кожному 10-му кадрі.

Кількісні характеристики якості ведення для алгоритмів Лукаса – Канаде, TLD та гістограмного підходу наведено в Табл. 1 (відео №1) та в Табл. 2 (відео №2). Алгоритми в таблицях пронумеровані так: A1 – алгоритм Лукаса – Канаде; A2 – алгоритм TLD; A3 – гістограмний метод.

Таблиця 1. Кількісні характеристики аналізу якості алгоритмів (відео №1: 1920x1080, 30 кадрів/сек, низький рівень шуму; площа розміченої області 2205 пікселів)

	\bar{d}	S_d	Швидкість, кадрів/сек
«довгострокове» ведення			
A1	2,71	1,70	28,02
A2	3,96	1,81	2,59
A3	5,31	2,30	6,49
«короткострокове» ведення			
A1	4,22	2,53	29,15
A2	8,83	3,44	6,09
A3	5,88	3,05	4,77

Таблиця 2. Кількісні характеристики аналізу якості алгоритмів (відео №2: 540x360, 30 кадрів/сек, значний рівень шуму; площа розміченої області 510 пікселів)

	\bar{d}	S_d	Швидкість, кадрів/сек
«довгострокове» ведення			
A1	9,87	6,16	30,00
A2	65,09	9,55	5,81
A3	6,70	3,51	21,04
«короткострокове» ведення			
A1	4,31	2,16	30,00
A2	58,26	12,98	8,54
A3	5,80	2,52	26,32

Проаналізувавши вищенаведені таблиці, можна вивести такі закономірності:

1. При умові низького рівня шуму (відео №1) якість довгострокового ведення є вищою. Для алгоритму TLD цей факт пов'язаний із більш якісним онлайн-навчанням на менш зашумлених даних. Для гістограмного методу – з отриманням менш зашумлених гістограм, які в цьому алгоритмі виступають у ролі дескрипторів. Імовірна причина для алгоритму Лукаса–Канаде – особливості розмітки тестових відео [7].

2. При умові високого рівня шуму (відео №2) якість короткострокового ведення є вищою.

3. Алгоритм TLD (у «класичному» варіанті; див. опис алгоритма, наведений вище) на зашумленому відео видає результат із середньою відстанню до розміченої області 60 пікселів при лінійних розмірах розміченої області 35x20 та розмірах кадру 540x360, тобто якість роботи для даного випадку є неприйнятною (рис. 1; тонкий прямокутник – розмітка, товстий – результат роботи алгоритму TLD).

4. Проаналізувавши середньоквадратичні відхилення, можна побачити, що:

4.1) на алгоритм Лукаса–Канаде шум впливає значною мірою; якщо рівень шуму низький, даний алгоритм ведення є найефективнішим із розглянутих;

4.2) алгоритм TLD має менше відхилення при «довгостроковому» веденні, але разом із цим падає швидкість роботи. Це є наслідком онлайн-навчання;

4.3) якість роботи гістограмного метода пропорційна площі розміченої області.

5. Гістограмний метод є найбільш робастним до шуму з розглянутих вище методів. Недоліком даного метода є зменшення швидкості його роботи пропорційно площі розміченої області (впливає з аналізу опису метода).

6. Алгоритм Лукаса–Канаде є найшвидшим із розглянутих, але менш точним при «довгостроковому» веденні.

Крім вищенаведеного, необхідно відмітити, що алгоритм TLD знаходить навіть об'єкти, які деякий час не знаходилися в полі зору камери.



Рис. 1. Приклад роботи алгоритму TLD на зашумленому відео

Висновки

Особливістю алгоритму Лукаса–Канаде є той факт, що він не враховує особливостей всієї області ведення, а лише деяку, як правило невелику, кількість опорних точок всередині неї. Негативними наслідками є відсутність можливості онлайн-навчання та висока імовірність накопичення помилки, позитивними – висока швидкість роботи.

Особливістю алгоритму TLD є можливість онлайн-навчання і, як наслідки, підвищення якості за рахунок більш точного знаходження об'єкта та можливість знаходити подібні об'єкти та об'єкти, які деякий час були відсутні у полі зору камери. Недоліки – низька швидкість роботи та некоректне навчання на зашумлених відео.

Гістограмний метод є відносно робастним до шуму і дає відносно якісні результати ведення об'єкта, але швидкість його роботи зменшується пропорційно площі розміченої області.

Подальші дослідження можуть бути направлені на зменшення вказаних недоліків за умови збереження наявних переваг.

Список літератури

1. Sandeep Kumar Patel et.al Moving Object Tracking Techniques: A Critical Review [електронний ресурс] / Sandeep Kumar Patel // Indian Journal of Computer Science and Engineering – 2013. – vol. 4 no. 2. – Режим доступу: <http://www.ijcse.com/docs/INDJCSE13-04-02-046.pdf>

2. Оптичний потік — Вікіпедія [електронний ресурс] // Режим доступу: https://uk.wikipedia.org/wiki/Оптичний_потік

3. Bruce D. Lucas, Takeo Kanade An Iterative Image Registration Technique with an Application to Stereo Vision [електронний ресурс] / Bruce D. Lucas // International Joint Conference on Artificial Intelligence – 1981. – pages 674–679. – Режим доступу: http://www.ces.clemson.edu/~stb/klt/lucas_bruce_d_1981_1.pdf

4. Алгоритм Лукаса – Канаде — Вікіпедія [електронний ресурс] // Режим доступу:

https://uk.wikipedia.org/wiki/Алгоритм_Лукаса_—_Канаде

5. Zdenek Kalal, Krystian Mikolajczyk, and Jiri Matas Tracking-Learning-Detection [електронний ресурс] / Zdenek Kalal // IEEE Transactions on Pattern Analysis and Machine Intelligence – 2010. – vol. 6 no. 1. – Режим доступу: http://kahlan.eps.surrey.ac.uk/featurespace/tld/Publications/2011_tpami

6. Rogatyuk A.A., Pristavka P.O. The Mathematical Foundation of Foreign Object Recognition in the Video From Unmanned Aircraft / Rogatyuk A.A. // Proceeding of the National Aviation University. – 2015 No. 3 (64). – P. 133–139.

7. Чирков А.В. Використання сплайнів при розмітці відео для тестування якості алгоритмів пошуку і ведення об'єктів / Чирков А.В. // XVI Міжнародна науково-практична конференція молодих учених і студентів "Політ. Сучасні проблеми науки". – Київ: НАУ – 2016.

Статтю подано до редакції 07.02.2016