

## Методи і средства обробки сигналів і зображень

УДК 004.932.2

**А.Ю. Варфоломєєв**, канд. техн. наук, **О.М. Лисенко**, д.-р. техн. наук

Національний технічний університет України «Київський політехнічний інститут»,  
вул. Політехнічна, 16/9, м. Київ, 03056, Україна.

### Визначення масштабу та повороту об'єкта шляхом кластеризації в методах відслідковування на основі обчислення оптичного потоку

*Запропоновано процедури оцінювання масштабу та повороту об'єкта за допомогою кластеризації, орієнтовані на застосування в методах відслідковування, що базуються на обчисленні розрідженого оптичного потоку. На основі зазначених процедур розроблено метод відслідковування під назвою FCT (Flow Clustering Tracker) та показано можливість забезпечення ним компромісного рівня надійності та точності відслідковування у порівнянні з існуючими рішеннями. Бібл. 8, рис. 5, табл. 2.*

**Ключові слова:** відслідковування; масштаб та поворот об'єкта; оптичний потік; кластеризація.

#### Вступ

Автоматичне візуальне відслідковування об'єктів знаходить широке застосування у робототехніці, медицині, кінематографії, військовій галузі, системах безпеки та спостереження тощо. Більшість із зазначених сфер передбачають використання вбудованих апаратних платформ, що накладає обмеження на обчислювальну складність алгоритмів, які мають на них реалізовуватись. Як наслідок, далеко не всі існуючі на сьогодні методи та алгоритми відслідковування можуть бути застосовані на практиці.

Серед наявних рішень можна виділити групу методів, в основі яких лежить оцінювання розрідженого оптичного потоку. Ці методи володіють помірною надійністю, але високою швидкістю [7], що робить доцільним їх застосування у вбудованих системах. Одним із найбільш ранніх представників даної групи методів є так званий метод медіанного потоку [4], який дозволяє окрім зсуву об'єкта також оцінювати його масштаб. В той же час він має ряд недоліків, серед яких наявність дрейфу області виявлення при повільному русі об'єкта та необхідність на кожному кадрі двічі обчислювати оптичний потік. З метою усунення зазначених

вад було розроблено модифікацію методу, відому як Flock of Trackers (FoT) [8]. Дана модифікація потребує обчислення оптичного потоку один раз на кадр, дозволяючи при цьому підвищити надійність відслідковування. Більш того, за рахунок використання в методі FoT алгоритму RANSAC (Random sample consensus) [1] стало можливим оцінювати не тільки масштаб об'єкта, але і його поворот. Тим не менш, нещодавно було запропоновано ще один підхід, який отримав назву Best Displacement Flow (BDF) [5]. Даний метод оснований на процедурі кластеризації векторів зміщення оптичного потоку і згідно тестам [7] володіє найбільшою надійністю серед розглянутих рішень. В той же час метод BDF на відміну від методів медіанного потоку та FoT дозволяє визначати лише переміщення об'єкта [5].

Мета даної роботи полягає у розширенні функціональності методу BDF шляхом розробки вдосконалень, які дозволили б додатково оцінювати масштаб та поворот об'єкта.

Нижче детальніше розглянуто методи медіанного потоку, FoT та BDF, наведено запропоновані вдосконалення методу BDF, які дозволяють оцінювати масштаб та поворот об'єкта, а також результати експериментальних досліджень на основі відео з тестової бази [7].

#### Короткі відомості про методи медіанного потоку, FoT та BDF

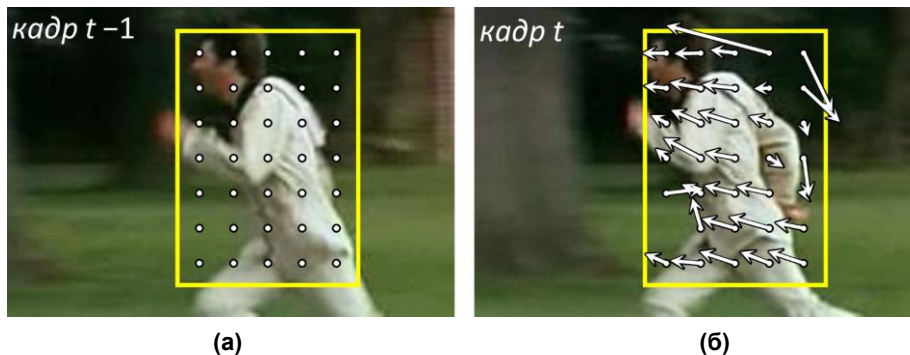
Для кращого розуміння запропонованих удосконалень, а також для повноти викладення коротко наведемо суть методів медіанного потоку [4], FoT [8] та BDF [5].

У всіх цих методах відслідковування об'єкта здійснюється шляхом аналізу зміщення окремих його точок між суміжними кадрами відеопослідовності. Обчислення зміщення виконується за допомогою алгоритмів оцінювання оптичного потоку: в методах медіанного потоку та FoT використовується пірамідальний алго-

ритм Лукаса-Канаде [2], а в BDF – блокове співставлення [5]. Точки, для яких обчислюється оптичний потік, обираються безпосередньо на об'єкті по розрідженій сітці (рис. 1, а).

Оптичний потік знаходить рух кожної з точок сітки (рис. 1, б). При цьому, як видно з рис. 1, результат по деяким точкам може виявитись невірним та не узгоджуватись із реальним рухом об'єкта. Причинами цього можуть бути так звана проблема апертури [3], випадкове потрапляння

точки на фоновий об'єкт, наявність перешкод тощо. Таким чином, щоб знайти дійсне переміщення об'єкта необхідно застосувати деяку процедуру, суть якої зводиться до відкидання точок, для яких зміщення знайдено некоректно та об'єднання результатів відстеження по тим точкам, що дали вірний результат. В залежності від того, за якою конкретно процедурою це виконується і відрізняються методи, що розглядаються.



**Рис. 1.** Точки об'єкта, за допомогою яких обчислюється зміщення об'єкта: а – розріджена сітка точок, для якої оцінюється оптичний потік; б – зміщення точок знайдені між суміжними кадрами

Згідно методу медіанного потоку оптичний потік обчислюється для сітки 10×10 точок [4]. При цьому точки, для яких зміщення знайдено некоректно, відфільтровуються на основі двох оцінок. Перша з них передбачає обчислення нормалізованої кореляції між околами відповідних точок на минулому та поточному кадрах. Друга називається помилкою прямого-зворотного відстеження (FB-помилкою) та обчислюється наступним чином:

$$E_{FB} = \| \mathbf{x} - \mathbf{x}' \| = \sqrt{(x - x')^2 + (y - y')^2},$$

де  $\mathbf{x} = (x, y)$  – координати точки на минулому кадрі;  $\mathbf{x}' = (x', y')$  – координати цієї ж точки, знайдені за допомогою обчислення оптичного потоку від минулого до поточного, а потім назад – від поточного до минулого кадрів.

Таким чином, FB-помилка є евклідовою відстанню між точками  $\mathbf{x}$  та  $\mathbf{x}'$ . На основі двох наведених оцінок відкидається 50 % найменш надійних точок, а за допомогою тих, що залишились, оцінюється зміщення та масштаб об'єкта.

Зміщення об'єкта відшукується як медіанне значення зсуву точок по кожній з координат:

$$d_x = \text{med}(\{\hat{x}_i - x_i\}),$$

$$d_y = \text{med}(\{\hat{y}_i - y_i\}),$$

де  $\text{med}(\cdot)$  – тут і надалі позначає статистичну медіану;  $\hat{\mathbf{x}}_i = (\hat{x}_i, \hat{y}_i)$  та  $\mathbf{x}_i = (x_i, y_i)$  – координати  $i$ -ої точки на поточному та минулому кадрах відповідно;  $\{\cdot\}$  – позначає множину різниць координат.

Масштаб об'єкта оцінюється наступним чином: беруться всі можливі пари точок на минулому та відповідні їм пари на поточному кадрі, обчислюються відстані між всіма парами точок та визначаються відношення між відстанями відповідних пар. У якості величини зміни масштабу  $\Delta s$  обирається медіанне значення усіх знайдених відношень:

$$\Delta s = \text{med} \left( \left\{ \frac{\| \hat{\mathbf{x}}_i - \hat{\mathbf{x}}_j \|}{\| \mathbf{x}_i - \mathbf{x}_j \|} \right\} \right),$$

де  $i$  та  $j$  – визначають всі можливі унікальні пари точок;  $\| \mathbf{x}_i - \mathbf{x}_j \|$  – позначає евклідову відстань між точками  $\mathbf{x}_i$  та  $\mathbf{x}_j$ ;  $\{\cdot\}$  – позначає множину всіх можливих відношень відстаней.

В даному методі на кожному кадрі генерується нова розріджена сітка точок і процедура відслідковування повторюється.

Розвитком розглянутого вище підходу є метод FoT [8]. У ньому для відкидання ненадійних точок використовуються дві процедури. Перша з них передбачає перевірку узгодженості

руху сусідніх точок на основі перетворення подібності [8]. Суть цієї перевірки полягає в наступному – фіксується деяка точка  $\mathbf{x}_i$  та розглядаються 4 її найближчі сусіди. Серед цих сусідів обираються пари точок  $\mathbf{x}_j$  та  $\mathbf{x}_k$ , для яких оцінюються матриці перетворення  $\mathbf{T}_{jk}$ , що переводять  $\mathbf{x}_j$  та  $\mathbf{x}_k$  у нові положення  $\hat{\mathbf{x}}_j$  та  $\hat{\mathbf{x}}_k$ , отримані після обчислення оптичного потоку. Далі перевіряється узгодженість руху точки  $\mathbf{x}_i$  з її сусідами за допомогою оцінки  $S_i^N$  (рис. 2):

$$S_i^N = \frac{1}{Z} \sum_{\substack{j,k \in N_i \\ j \neq k}} \mathbf{1}(\|\mathbf{T}_{jk}\mathbf{x}_i - \hat{\mathbf{x}}_i\|^2 < \varepsilon_N),$$

де  $Z$  – нормуючий коефіцієнт, який обирається таким, щоб  $0 \leq S_i^N \leq 1$ ;  $N_i$  – індекси сусідніх до  $\mathbf{x}_i$  точок;  $\mathbf{1}(\cdot)$  – функція, яка приймає значення 1 тоді, коли виконується умова в її аргументі та дорівнює 0 – в іншому випадку;  $\varepsilon_N$  – допустиме відхилення (у роботі [8]  $\varepsilon_N = 2$ ).

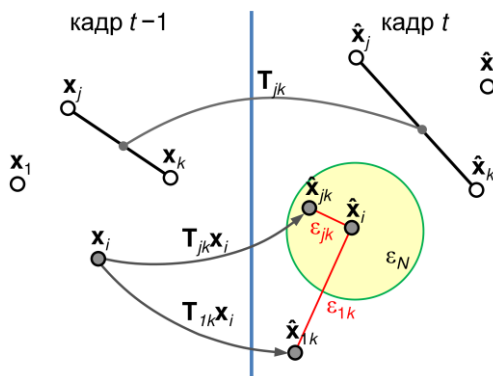


Рис. 2. Перевірка узгодженості руху точки із її сусідами у методі FoT

Якщо оцінка  $S_i^N$  виявиться більшою за деякий поріг  $\theta_N$ , то зміщення точки  $\mathbf{x}_i$  у нове положення  $\hat{\mathbf{x}}_i$  на поточному кадрі, знайдене оптичним потоком, вважається узгодженим із сусідніми точками (у роботі [8] поріг  $\theta_N = 1/6$ ). Точки, що рухаються не узгоджено, вважаються ненадійними і відкидаються.

Друга процедура визначення надійних точок основана на наступному припущенні: точки, які відслідковувались коректно на попередніх кадрах найбільш імовірно мають давати вірний результат і на поточному кадрі [8]. Таким чином, виникає необхідність передбачення стану точки в поточний момент на основі попередніх спостережень за нею. Цю задачу природно вирішу-

вати за допомогою марківського ланцюга. З огляду на це в методі використовується простий стохастичний автомат з двома станами: 1) точка є ненадійною, її називають «викидом» (outlier); 2) точка є надійною, її називають «не викидом» (inlier). Перехід між станами виконується на основі визначення відносних частот перебування точки в стані «викид»/«не викид» протягом часу відслідковування. Детальніше з даною процедурою можна ознайомитись в роботі [8].

Згідно методу FoT після застосування процедур виявлення надійних точок нове положення об'єкту обчислюється за допомогою алгоритму RANSAC [1]. Цей алгоритм випадковим чином формує набори точок, за якими буде імовірна модель перетворення координат об'єкта (афінні, проєктивні тощо). Далі він перевіряє, наскільки кожна побудована модель відповідає переміщенню точок об'єкта, що не приймали участь в побудові даної моделі та у якості дійсної трансформації об'єкта приймає ту модель, яка дає найменше відхилення. За рахунок використання алгоритму RANSAC метод FoT дозволяє оцінювати як переміщення та масштаб об'єкта, так і його поворот або навіть проєктивне перетворення.

У методі FoT точки, які визнаються надійними, продовжують відслідковуватись на подальших кадрах, тоді як викиди повертаються у відповідні початкові положення на сітці.

Згідно методу BDF [5] нове положення об'єкта оцінюється за допомогою простої процедури кластеризації, яка розподіляє точки на групи так, що в кожену групу потрапляють точки з однаковим (або приблизно однаковим) зміщенням. Найбільша група, очевидно, відповідає зміщенню, яке було знайдене оптичним потоком у більшості випадків, тому його вважають найбільш імовірним і називають *вектором найкращого зміщення*  $\beta$  (рис. 3, а [5]). Положення об'єкту задається як  $\hat{\mathbf{O}}_{bb} = \mathbf{O}_{bb} + \beta$ , де  $\mathbf{O}_{bb}$  та  $\hat{\mathbf{O}}_{bb}$  – це центри прямокутників, що задають положення об'єкта на минулому та поточному кадрах відповідно.

Для фільтрації ненадійних точок подібно до того, як це реалізовано у методі FoT, використовується поняття стану  $i$ -ої точки:

$$S_i = \begin{cases} \text{викид}, & \text{якщо } \|d_i - \beta\|_{\infty} > \delta_s, \\ \text{невикид}, & \text{інакше,} \end{cases}$$

де  $d_i$  – зміщення  $i$ -ої точки;  $\|\cdot\|_{\infty}$  – нескінченна норма;  $\delta_s$  – максимально допустиме відхилення зміщення від вектора  $\beta$ .

Точки, які визнаються викидами, відкидаються і не приймають участі в оцінці зміщення об'єкта на подальших кадрах. При цьому, якщо кількість точок, що вважаються надійними (не викидами), впаде нижче деякого порогу  $\delta_N$ , то відбувається їх відновлення у початкових положеннях на сітці. Тим не менш відновлені точки одразу не приймають участь в оцінці зміщення об'єкта, а переводяться в проміжний стан, з якого вони можуть стати надійними, якщо впродовж

$\delta_t$  кадрів даватимуть зміщення, що за відхиленням від  $\beta$  не перевищуватиме  $\delta_s$  або знову перейти в стан ненадійних (викидів). У роботі [5] наведена процедура відновлення положень точок називають *консенсусною переініціалізацією* (рис. 3, б [5]), її параметрами є  $\delta_s = 7$ ,  $\delta_t = 3$ , а  $\delta_N = 25\%$  від загальної кількості точок, що мають відслідковуватись.

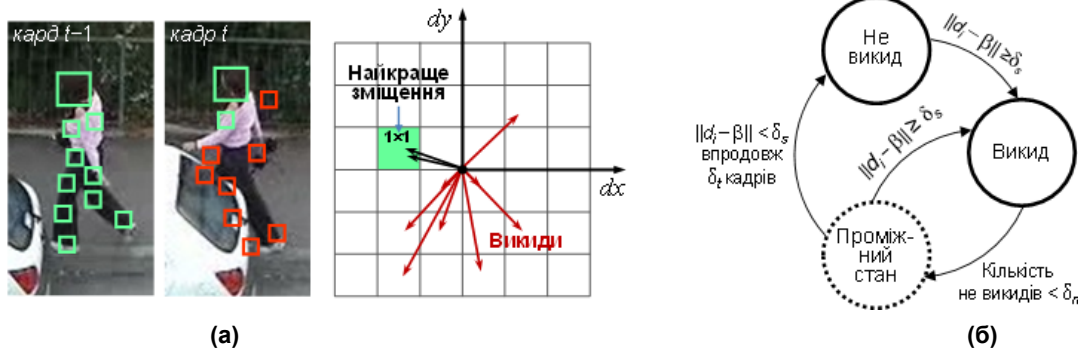


Рис. 3. Метод BDF: а) – кластеризація зміщення точок; б) – процедура консенсусної переініціалізації

У методі BDF, на відміну від двох розглянутих вище підходів, для обчислення оптичного потоку використовується блокове співставлення, в якому для підвищення надійності одночасно використовуються два розміри блоків  $7 \times 7$  та  $13 \times 13$  [5].

**Оцінювання масштабу та повороту об'єкта за допомогою процедури кластеризації**

Як вже зазначалося, під час тестування метод BDF показав більш високу надійність відслідковування, ніж методи медіанного потоку та FoT [7]. В той же час метод BDF не передбачає можливості оцінки масштабу та повороту об'єкта, що може бути корисним у ряді випадків. З метою подолання цього обмеження пропону-

ються дві процедури, які також основані на кластеризації.

Зокрема, для визначення масштабу об'єкта проводиться розрахунок відстаней між всіма парами точок на минулому та поточному кадрах. Після цього між знайденими відстанями відповідних пар обчислюються різниці (рис. 4, а):

$$\Delta d_{i,j} = \|\hat{x}_i - \hat{x}_j\| - \|x_i - x_j\|, \quad (1)$$

де  $\hat{x}_i = (\hat{x}_i, \hat{y}_i)$  та  $x_i = (x_i, y_i)$  – координати  $i$ -ої точки на поточному та минулому кадрах відповідно;  $i$  та  $j$  – визначають всі можливі унікальні пари точок;  $\|x_i - x_j\|$  – позначає евклідову відстань між точками  $x_i$  та  $x_j$ .

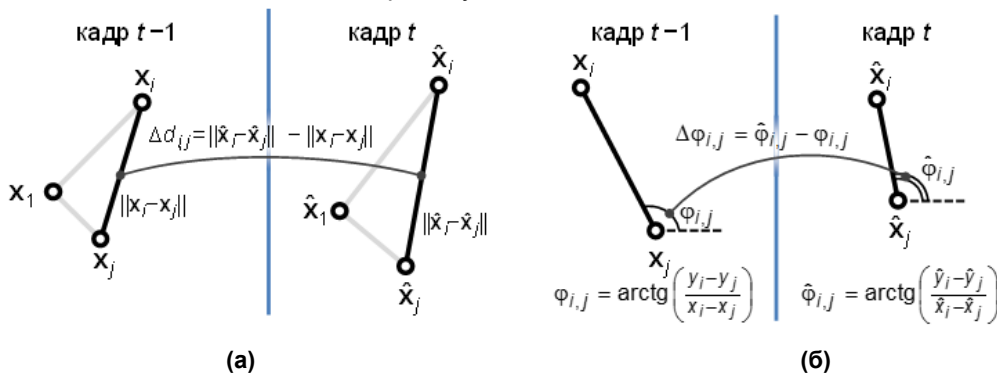


Рис. 4. Обчислення різниць відстаней – а та кутів нахилу – б між відповідними парами точок на попередньому та наступному кадрах. Різниці показані тільки для пари точок  $i$  та  $j$ , проте в процедурах, що розглядаються, вони повинні обчислюватись для всіх можливих унікальних пар точок

Знайдені різниці розподіляються на групи, для чого застосовується процедура кластеризації, така сама як в методі BDF для оцінки зсуву. Оскільки ця процедура полягає у пошуку однакових значень, а відстані  $\Delta d_{i,j}$  є дійсними числами, які можуть незначно відрізнятися, для запобігання надмірної кластеризації у якості її критерію пропонується використовувати абсолютне значення  $\Delta d_{i,j}$ , помножене на коефіцієнт квантування та округлене до найближчого цілого:

$$\Delta_q d_{i,j} = [\Delta d_{i,j} \cdot q_d], \quad (2)$$

де оператор  $[\cdot]$  – позначає округлення до найближчого цілого значення;  $q_d$  – коефіцієнт квантування, який визначає «грубість» округлення, впливаючи на «точність» кластеризації. В експериментах  $q_d$  було прийнято рівним 1.

Після кластеризації величин  $\Delta_q d_{i,j}$ , обчислених за допомогою (2), визначається найбільший кластер  $K$  та розраховується середнє арифметичне відношень відстаней між парами точок, що в нього ввійшли:

$$\Delta s = \frac{1}{|K|} \sum_{(i,j) \in K} \frac{\|\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_j\|}{\|\mathbf{x}_i - \mathbf{x}_j\|}, \quad (3)$$

де  $|K|$  – кількість елементів в кластері  $K$ .

Обчислена величина  $\Delta s$  вважається зміною масштабу об'єкта між минулим і поточним кадрами.

Оцінювання повороту об'єкта виконується аналогічно до визначення масштабу, тільки замість різниць відстаней тут кластеризуються різниці кутів нахилу між прямими, що проходять через всі відповідні пари точок минулого та поточного кадрів (рис. 4, б):

$$\Delta_q \varphi_{i,j} = \left[ \left( \arctg \left( \frac{\hat{y}_i - \hat{y}_j}{\hat{x}_i - \hat{x}_j} \right) - \arctg \left( \frac{y_i - y_j}{x_i - x_j} \right) \right) \cdot q_\varphi \right], \quad (4)$$

де  $\hat{\mathbf{x}}_i = (\hat{x}_i, \hat{y}_i)$  та  $\mathbf{x}_i = (x_i, y_i)$  – як і раніше координати  $i$ -ї точки на поточному та минулому кадрах відповідно;  $i$  та  $j$  – визначають всі можливі унікальні пари точок; оператор  $[\cdot]$  – позначає округлення до найближчого цілого значення;  $q_\varphi$  – коефіцієнт квантування, що впливає на точність кластеризації.

При проведенні експериментальних досліджень використовувався  $q_\varphi = 1$  за умови, що функція  $\arctg(\cdot)$  дає результат у градусах (кластеризація повороту виконується з точністю до  $1^\circ$ ).

Після кластеризації величини  $\Delta_q \varphi_{i,j}$  обирається найбільший кластер  $K$  і для не округлених різниць кутів нахилу розраховується середнє арифметичне значення, яке вважається поворотом об'єкта між минулим і поточним кадрами:

$$\Delta \varphi = \frac{1}{|K|} \sum_{(i,j) \in K} \left( \arctg \left( \frac{\hat{y}_i - \hat{y}_j}{\hat{x}_i - \hat{x}_j} \right) - \arctg \left( \frac{y_i - y_j}{x_i - x_j} \right) \right), \quad (5)$$

де  $|K|$  – кількість елементів в кластері  $K$ .

Зауважимо, що наведені вище процедури мають дещо спільне з алгоритмом RANSAC – вони фактично здійснюють пошук моделі руху об'єкта, яка відповідає зміні координат більшості його точок. В той же час запропонований підхід має одну важливу відмінність – він дозволяє оцінювати параметри руху об'єкта (зміщення, зміну масштабу та повороту) незалежно один від одного. Це робить його менш чутливим до відхилень окремих параметрів, а також забезпечує вищу гнучкість, дозволяючи використовувати для кожної процедури оцінки окремі налаштування. Як буде показано в наступному розділі, система, що застосовує запропонований підхід, володіє дещо вищою надійністю, ніж метод FoT, оснований на алгоритмі RANSAC.

## Результати експериментальних досліджень

Експериментальні дослідження проводились на програмній реалізації модифікованого варіанту методу BDF, названого FCT (Flow Clustering Tracker). У даній реалізації оптичний потік оцінювався за допомогою пірамідального методу Лукаса-Канаде [2]. При цьому для підвищення надійності роботи з об'єктами малого розміру його обчислення здійснювалось з урахуванням різної кількості рівнів піраміди: для однієї половини точок оптичний потік рахувався на 3-х рівнях піраміди, а для іншої – на 5-и. Початкова ініціалізація положень точок виконувалась із застосуванням методу FAST [6] таким чином, щоб знайдені за допомогою даного детектора ключові точки приблизно розташовувались у вузлах регулярної сітки розміром  $10 \times 10$  (рис. 1). Оцінювання зміщення проводилось за допомогою процедури кластеризації, як і в методі BDF. Точність оцінки зміщення кожної точки при цьому округлялась до 1 пікселя. Вектор найкращого зміщення  $\beta$  оцінювався як середнє арифметичне неокруглених координат змішень точок, які потрапили у найбільший кластер. Визначення міжкадрової зміни масштабу  $\Delta s$  та повороту  $\Delta \varphi$  об'єкта виконувалось за процеду-



рами, описаними в попередньому розділі згідно виразів (1)-(5). Виявлення ненадійних точок проводилось на основі відхилення їх зміщень від вектора найкращого зсуву  $\beta$ . Зокрема, точка вважалася ненадійною (викидом), якщо виконувалась умова:

$$\begin{cases} \|d_i - \beta\| > k \cdot \|\beta\| \\ \|\beta\| > \rho_{OF} \end{cases},$$

де  $d_i$  – зміщення  $i$ -ої точки;  $\|\beta\|$  – довжина вектора  $\beta$ ;  $k$  – коефіцієнт пропорційності;  $\rho_{OF}$  – точність обчислення оптичного потоку.

Параметри  $k$  та  $\rho_{OF}$  були підібрані емпірично і прийняті рівними 1,75 та 0,25 відповідно. Якщо кількість точок, що була визнана ненадійними, перевищувала 25% від загальної кількості, то здійснювалась нова ініціалізація їх положень за сіткою без використання методу FAST.

Експериментальні дослідження проводились в рамках VOT 2015 і також представлені у [7] системою відслідковування під назвою FCT. VOT 2015 – це еталонний тест, що містить 60 складних короткотривалих відеопослідовностей, на яких має відслідковуватись єдиний об'єкт. Для кожної відеопослідовності еталонні положення об'єктів є відомими (визначені вручну). Важливим елементом VOT є підсистема тестування, яка дозволяє проводити автоматизовану оцінку якості роботи методів відслідковування. Нею використовуються наступні критерії: *точність A* – величина, яка характеризує перекриття еталонного і знайденого системою відслідковування положень об'єкта; *надійність R* – усереднена кількість зривів відслідковування на тестових відеопослідовностях; *міра*, яка об'єднує

обидва ці показники – *очікуване усереднене перекриття  $\hat{\Phi}$*  еталонного й оціненого системою, що тестується, положень об'єкта; а також відносна *швидкодія*, яка виражається в так званих еквівалентних операціях фільтрації [7].

Відповідно до результатів [7] у табл. 1 наведено порівняльні характеристики методів FoT, BDF та FCT за критеріями точності, надійності та очікуваного усередненого перекриття.

Таблиця 1. Показники точності A, усередненої кількості зривів R та очікуваного усередненого перекриття  $\hat{\Phi}$  для розглянутих методів

Метод	A↑	R↓	$\hat{\Phi}$ ↑
FoT	0,43	4,36	0,14
BDF	0,4	3,11	0,15
FCT	0,43	3,34	0,15

Стрілки біля критеріїв позначають, що система володіє кращими характеристиками, якщо: критерій вище – «↑», критерій нижче – «↓». Порівняння з методом медіанного потоку не проводилось, оскільки відомо, що він дає гірші результати, ніж метод FoT [4]. Як видно з табл. 1, метод FCT має вищу точність у порівнянні з методом BDF, але дещо поступається йому в надійності. В той же час метод FCT є більш надійним, ніж метод FoT, забезпечуючи при цьому аналогічну точність. Очікуване усереднене перекриття є однаковим для методів FCT та BDF і переважає відповідний показник для методу FoT.

На рис. 5, а) наведено графік усередненої кількості зривів на 100 кадрів для кожного з методів, що порівнюються [7].



(а)



(б)

Рис. 5. Усереднена кількість зривів на 100 кадрів в залежності від атрибутів відеопослідовностей – а); приклади відслідковування об'єктів методом FCT – б)

З рис. 5 можна бачити, що FCT забезпечує нижчу надійність лише при зміні освітлення та перекритті об'єкта. Це найбільш імовірно обумовлено використанням оптичного потоку на основі метода Лукаса-Канаде, який є чутливим до змін яскравості та не накопичує інформацію про зовнішній вигляд об'єкта. Також причиною зниження надійності може бути відсутність у методі FCT процедур відновлення/відкидання ненадійних точок, подібних до консенсусної переініціалізації або марківського ланцюга, що використовуються у методах BDF та FoT відповідно.

На рис. 5, б) наведено приклади відслідковування об'єктів методом FCT на деяких відеопослідовностях бази VOT 2015.

Показники швидкодії метода FCT, отримані на різних апаратних платформах, зведено до табл. 2. Метод FCT при цьому було реалізовано на мові C++ із застосуванням бібліотеки OpenCV версії 2.4.10.

Табл. 2. Швидкодія метода FCT на різних апаратних платформах та відео з різною роздільною здатністю

Платформа	Середній час обробки кадру, мс	
	768×576	320×240
Intel Core i7 (3,5 ГГц)	2,1	1,1
Intel Core Duo (1,86 ГГц)	8,4	4,1
ARM Cortex A8 (1 ГГц)	99,5	22,5

## Висновки

В роботі запропоновано нові процедури оцінювання масштабу та повороту об'єкта в методах відслідковування, основаних на обчисленні оптичного потоку. Суть нововведень зводиться до кластеризації різниць довжин та кутів нахилу відрізків, що сполучають відповідні пари точок на суміжних кадрах відеопослідовності (рис. 4). На основі зазначених процедур розроблено метод відслідковування під назвою FCT (Flow Clustering Tracker), який, як було показано за допомогою відкритого тесту систем відслідковування VOT 2015, дозволяє досягти компромісного рівня надійності та точності відслідковування в порівнянні з існуючими рішеннями (FoT та BDF). Розроблені процедури та метод FCT також володіють достатньо високою швидкодією, що дозволяє застосовувати їх у вбудованих та портативних системах відслідковування.

Напрямок подальших досліджень може стати підвищення надійності методу FCT за рахунок розробки та застосування у ньому гібридних процедур обчислення оптичного потоку, основаних на одночасному застосуванні методів Лукаса-Канаде та блокового співставлення.

## Список використаних джерел

1. *Форсайт Д.* Компьютерное зрение. Современный подход.: Пер. с англ. / Д. Форсайт,

- Ж. Понс – М.: Издательский дом «Вильямс», 2004. – 928 с.
2. *Boguet J.-Y.* Pyramidal Implementation of the Lucas-Kanade Feature Tracker: Description of the algorithm. Tech. report, Intel Corporation. Microprocessor Research Labs / [J.-Y. Boguet] – 2002. – 9 p.
3. *Bradski G.* Learning OpenCV / G. Bradski, A. Kaehler. – Sebastopol: O'Reilly Media, Inc., 2008. – 557 p.
4. *Kalal Z.* Forward-Backward Error: Automatic Detection of Tracking Failures / Z. Kalal, K. Mikolajczyk, J. Matas // International Conference on Pattern Recognition, Istanbul. – 2010. – P. 2756–2759.
5. *Maresca M.E.* Clustering Local Motion Estimates for Robust and Efficient Object Tracking / M.E. Maresca, A. Petrosino // Lecture Notes in Computer Science. – Vol. 8926. – 2015. – P. 244–253.
6. *Rosten E.* Machine learning for high-speed corner detection / E. Rosten, T. Drummond // European Conference on Computer Vision (ECCV), Graz, 2006 p. : Proceedings. – 2006. – P. 430–443.
7. The Visual Object Tracking VOT2015 challenge results / [ M. Kristan., J. Matas, A. Leonardis et al ] // IEEE International Conference on Computer Vision Workshop (ICCVW), Santiago. – 2015. P. 564–586.
8. *Vojtř T.* Increasing Robustness of the Flock of Trackers / T. Vojtř, J. Matas // Research Reports of CMP, Czech Technical University in Prague. – 2012. – No. 14. – 20 p.

Поступила в редакцию 25 мая 2016 г.

УДК 004.932.2

**А.Ю. Варфоломеев**, канд. техн. наук, **А.Н. Лысенко**, д. – р. техн. наук

Национальный технический университет Украины «Киевский политехнический институт»,  
ул. Политехническая, 16/9, г. Киев, 03056, Украина.

## **Определение масштаба и поворота объекта путем кластеризации в методах отслеживания на основе вычисления оптического потока**

*Предложены процедуры оценки масштаба и поворота объекта с помощью кластеризации, ориентированные на применение в методах отслеживания, основанных на вычислении разреженного оптического потока. На основе указанных процедур разработан метод отслеживания под названием FCT (Flow Clustering Tracker) и показана его способность обеспечить компромиссный уровень надежности и точности отслеживания в сравнении с существующими решениями. Библи. 8, рис. 5, табл. 2.*

**Ключевые слова:** отслеживание; масштаб и поворот объекта; оптический поток; кластеризация.

UDC 004.932.2

**A. Varfolomieiev**, Ph.D., **O. Lysenko**, Dr.Sc.

National Technical University of Ukraine «Kyiv Polytechnic Institute»,  
Polytechnichna str., 16/9, Kyiv, 03056, Ukraine.

## **Object's scale and rotation estimation using the clustering in tracking methods based on the optic flow calculation**

*Procedures of scale and rotation estimation of an object using the clustering are proposed. The procedures are oriented to the application in tracking techniques based on sparse optic flow calculation. The new tracking method, which is called FCT (Flow Clustering Tracker) is developed using the mentioned procedures. The obtained results show its ability to achieve the compromise in robustness and accuracy with respect to the existing solutions. References 8, figures 5, tables. 2.*

**Keywords:** tracking; scale and rotation of an object; optic flow; clustering.



**References**

1. *Forsyth, D., & Ponce, J.* (2004). *Computer vision: A modern approach*. Moscow: Williams Publishing House. (Rus)
2. *Boguet, J.-Y.* (2002). *Pyramidal Implementation of the Lucas-Kanade Feature Tracker: Description of the algorithm*. (Tech. report, Intel Corporation. Microprocessor Research Labs).
3. *Bradski, G., & Kaehler, A.* (2008). *Learning OpenCV*. Sebastopol (CA): O'Reilly.
4. *Kalal, Z., Mikolajczyk, K., & Matas, J.* (2010). *Forward-Backward Error: Automatic Detection of Tracking Failures*. 2010 20th International Conference on Pattern Recognition, 2756-2759. doi:10.1109/icpr.2010.675.
5. *Maresca, M.E., & Petrosino, A.* (2015). *Clustering Local Motion Estimates for Robust and Efficient Object Tracking*. *Computer Vision - ECCV 2014 Workshops Lecture Notes in Computer Science*, 244-253. doi:10.1007/978-3-319-16181-5\_17.
6. *Rosten, E., & Drummond, T.* (2006). *Machine Learning for High-Speed Corner Detection*. *Computer Vision – ECCV 2006 Lecture Notes in Computer Science*, 430-443. doi:10.1007/11744023\_34.
7. *Kristan, M., Matas, J., Leonardis, A., Felsberg, M., Cehovin, L., Fernandez, G., Pflugfelder, R.* (2015). *The Visual Object Tracking VOT2015 Challenge Results*. 2015 IEEE International Conference on Computer Vision Workshop (ICCVW). doi:10.1109/iccvw.2015.79.
8. *Vojšíř, T. & Matas, J.* (2012). *Increasing Robustness of the Flock of Trackers* (Research Report of CMP No. 14., Czech Technical University in Prague).