

Міністерство освіти і науки України
Національний технічний університет України
Київський політехнічний інститут ім. Ігоря Сікорського
Навчально-науковий комплекс
Інститут прикладного системного аналізу

Магістерська дипломна робота на тему:
Система трекінгу об'єктів на відеоряді в
режимі реального часу

Виконав: студент VI курсу
Групи КА-51м
Лисенко Андрій
Науковий керівник:
к.т.н. доц. Дідковська М.В.

Актуальність роботи

- впровадження систем відеоспостереження в автоматизовані системи управління технологічними процесами
- ефективне використання каналів передачі даних за рахунок передачі технічної (аналітичної) метаінформації замість передачі зображення.

Постановка задачі

Мета роботи: Побудова системи трекінгу об'єктів на відеоряді в режимі реального часу.

Об'єкт дослідження: Система трекінгу об'єктів на відеоряді в режимі реального часу.

Предмет дослідження: методи трекінгу швидких та повільних об'єктів на відеоряді.

Постановка задачі

- Проаналізувати існуючі підходи трекінгу;
- Розробити алгоритм трекінгу об'єктів на відеоряді;
- Мінімізувати обчислювальну складність для пришвидшення обчислень;
- Реалізувати алгоритм в програмному продукті.

Предметна область

В рамках даної роботи введено обмеження одночасної наявності єдиного екземпляра шуканого об'єкта в кадрі.

Задачу трекінгу розбивають на:

- Розпізнавання об'єкта
- Передбачення положення об'єкта

Огляд існуючих підходів

Методи розпізнавання об'єктів:

- SIFT
висока обчислювальна складність
- SURF
менша точність
- PREDATOR
необхідність тренування перед/в процесі використання

Методи прогнозування положення:

- Моделювання сегмента траєкторії
необхідність великої кількості специфічних підготованих даних
- Фільтр Калмана
менша точність
- P- та N-експерти (PREDATOR)
ускладнена реалізація при використанні окремо від основного алгоритму

Оцінки якості

$$Q = \frac{N_{sd}}{N_{st}} \cdot \left(1 - \frac{N_{nd}}{N_{nt}} \right) \cdot 100\%$$

- N_{sd} – к-ть кадрів, які розпізнано правильно
- N_{st} – загальна кількість кадрів
- N_{nd} – к-ть кадрів, які розпізнано неправильно
- N_{nt} – загальна кількість кадрів, що не містять об'єкт

Порівняльна характеристика

- В результаті проведеного порівняння було отримано оцінки для розглянутих методів обчислення характеристик

Метод	Точність	Середня тривалість обробки кадру
SIFT	91.3%	0.06 с
SURF	87.4%	0.02 с
PREDATOR	94.5%	0.05 с

Порівняльна характеристика

- В результаті проведеного порівняння було отримано оцінки для розглянутих методів прогнозування положення

Метод	Середня помилка	Середня тривалість обробки кадру
Моделювання сегмента траєкторії	0.64 px	1.2 с
Фільтр Калмана	0.79 px	0.3 с
PREDATOR	0.74 px	0.72 с

SURF

- Гессіан - досягає екстремуму в точках максимального зміни градієнта яскравості, добре виділяє плями, кути і краї

$$H(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix}$$

$$\det(H_{approx}) = D_{xx}D_{yy} - (0.9D_{xy})^2$$

$$V = (\Sigma d_x, \Sigma d_y, \Sigma |d_x|, \Sigma |d_y|)$$

Фільтр Калмана

- Фільтр Калмана є ефективним рекурсивним фільтром, що може оцінити наступний стан динамічної системи, засновану на зашумленому вимірі.

$$X(k | k - 1) = AX(k - 1 | k - 1) + BU(k) + W(k)$$

$$P(k | k - 1) = AP(k - 1 | k - 1)A' + Q$$

$$X(k | k) = X(k | k - 1) + Kg(k)(HX(k) - HX(k | k - 1))$$

$$Kg(k) = \frac{P(k | k - 1)H'}{HP(k | k - 1)H' + R}$$

$$P(k | k) = (1 - Kg(k)H)P(k | k - 1)$$

Модифікований алгоритм

- У нашій системі немає елемента управління, тому $U(k) = 0$. Якщо немає локального руху, припускаємо, що наступна оцінка стану $X(k | k - 1)$ дорівнює попередньому оптимізованому стану $X(k - 1 | k - 1)$, а $A = 1$ та $H = 1$, також $Q = 1, R = 0.05$.

$$X(k|k - 1) = X(k - 1 | k - 1)$$

$$P(k|k - 1) = P(k - 1 | k - 1) + 1$$

$$X(k|k) = X(k|k - 1) + Kg(k)(HX(k) - X(k|k - 1))$$

$$Kg(k) = \frac{P(k | k - 1)}{P(k | k - 1) + R}$$

$$P(k | k) = (1 - Kg(k))P(k | k - 1)$$

Приклад роботи



Результати роботи

- Алгоритм було випробувано на 8 типах відеоматеріалів

Типи матеріалів	Точність розпізнавання	Середня похибка передбачення	Швидкодія
Тип 1	83.6%	0.84 рх	0.29 с
Тип 2	85.7%	0.81 рх	0.31 с
Тип 3	92.4%	0.72 рх	0.30 с
Тип 4	91.8%	0.702 рх	0.28 с
Тип 5	92.6%	0.701 рх	0.29 с
Тип 6	91.7%	0.73 рх	0.28 с
Тип 7	97.4%	0.31 рх	0.31 с
Тип 8	95.9%	0.58 рх	0.29 с
Середнє значення	91.4%	0.674 рх	0.29 с

Аналіз результатів

- запропонований метод розпізнавання розпізнає об'єкти в середньому краще ніж два з трьох розглянутих раніше методів поступаючись лише методу, який використовує засоби штучного інтелекту для класифікації зображень. Також запропонований метод передбачення положення показав себе загалом краще ніж два з трьох розглянутих методів, поступаючись лише методу, що базується на моделюванні сегмента траєкторії.
- оцінка якості складає 91.4%, що більше за 91.3% та 87.4%, оцінка точності складає 0.674 px, що менше за 0.74 px та 0.78 px

Висновки

Новизна

- Запропоновано алгоритм трекінгу об'єктів, що комбінує метод розпізнавання об'єктів SURF та метод прогнозування положення об'єкта базований на модифікованому фільтрі Калмана.

Практична цінність

- Створено систему, що розпізнає та прогнозує положення об'єкта в середньому краще ніж два з трьох методів розпізнавання та два з трьох методів прогнозування положення.

Публікації

- Порівняльний аналіз методів трекінгу об'єктів на відеоряді, матеріали конференції САІТ 2017, м.Київ, 22-25 травня 2017
- Порівняльний аналіз методів передбачення траєкторії об'єкта в задачі трекінгу об'єкта на відеоряді, матеріали міжнародної інтернет-конференції «Весняні наукові читання», м.Вінниця, 22 травня 2017

Можливості розвитку

- Використання засобів штучного інтелекту для класифікації (нейронні мережі, SVM);
- Розширення кількості використовуваних методів розпізнавання об'єктів;
- Розширення кількості методів прогнозування положення об'єкта;
- Розгляд випадку одночасної наявності в кадрі декількох екземплярів шуканого об'єкта;