

О ВАЖНОСТИ ОПТИМИЗАЦИИ ТОЧНОСТИ И ПРОИЗВОДИТЕЛЬНОСТИ АЛГОРИТМОВ ОТСЛЕЖИВАНИЯ ОБЪЕКТОВ В ВИДЕОПОТОКЕ ПО ПАРАМЕТРАМ

© 2019 г. А. Д. Егоров*

Национальный исследовательский ядерный университет "МИФИ", Москва, 115409, Россия

**e-mail: ADEgorov@mephi.ru*

Поступила в редакцию 15.01.2019 г.

После доработки 28.02.2019 г.

Принята к публикации 12.03.2019 г.

В статье рассматривается вопрос повышения эффективности работы алгоритмов отслеживания объектов на изображении за счет оптимизации по параметрам для данных алгоритмов. В отличие от большей части подобных исследований в целевой функции оптимизации учитывается не только точность работы алгоритмов отслеживания, но и их производительность. С целью составления алгоритма расчет целевой функции в работе приводится анализ существующего пространства ситуаций в работе рассматриваемых алгоритмов, показано, что в реальной работе существует 5 типов ситуаций: корректное отслеживание, ложное отслеживание, преждевременная потеря объекта, объект не найден вовремя, объект потерян и найден повторно. Правильный подбор параметров исходного алгоритма, на примере алгоритма отслеживания объектов по шаблону, и его модификаций позволяет добиться роста точности в 4.7 раза при росте производительности в 2.2 раза. При этом, показано, что отсутствие подбора параметров может на практике приводить к отказу от использования эффективных модификаций. В частности, модификация адаптации шаблона поиска в 90% случаев ухудшает работу алгоритма отслеживания по шаблону, а в 10% значительно улучшает его работу. Кроме того, в статье приводятся конкретные параметры алгоритма отслеживания по шаблону и его модификаций, которые целесообразно использовать в работе систем компьютерного зрения.

Ключевые слова: оптимизация параметров, отслеживание по шаблону в видеопотоке, компьютерное зрение, отслеживание объектов в видеопотоке, видеопоток

DOI: 10.1134/S2304487X19030052

ВВЕДЕНИЕ

Компьютерное зрение на текущий момент широко применяется в различных межмашинных и человеко-машинных системах [1]. Особо актуальными являются следующие задачи: автоматизация систем управления, например, для робототехнических систем или беспилотных систем [2]; обеспечения безопасности, например, дорожного движения [3] или периметра охраняемого объекта; обеспечение безопасности критических объектов [4, 5]; ускорение производства; поддержка принятия решения, например, при проведении хирургических операций [6] и т.д.

К задачам создания систем управления беспилотными автомобилями и иными подобными роботизированными системами предъявляются не только требования по точности, но и повышенные требования по автономности работы подобных систем [2]. То есть для таких систем необходимо кроме реализации самих алгоритмов проводить их максимальную оптимизацию.

Задача оптимизации не нова, ее решение для различных частных случаев приводится в ряде работ различных авторов [7–9]. Тем не менее, постановка задачи в данных работах не рассматривает вопросы параметрической оптимизации алгоритмов отслеживания объектов в видеопотоке, касаясь вскользь только параметрической оптимизации для задач поиска объектов на изображении. Потенциал подобной оптимизации достаточно высок, что показано на схожих работах, связанных с оптимизацией самих алгоритмов поиска объекта [10–12], ведь поиск объекта на изображении является предваряющей стадией отслеживания объектов.

В данной статье рассматривается влияние параметрической оптимизации на работу алгоритмов отслеживания, показывается, что параметрическая оптимизация является необходимым элементом реализации любого алгоритма отслеживания объектов в видеопотоке. Составными элементами данной работы является конструирование собственной метрики для сравне-

ния алгоритмов отслеживания объектов, а также проверка на практике результатов теоретической разработки.

ТЕОРИЯ

Существующие алгоритмы отслеживания объектов в видеопотоке можно разделить на несколько категорий по принципу, который лежит в основании их реализации [13]:

1. Сравнение с помощью математических методов представления предполагаемого объекта с его эталонным представлением, сохраненным в памяти [14, 15].
2. Сравнение представления предполагаемого объекта с регулярно модифицируемым хранимым в памяти представлением объекта [16, 17].
3. Сравнение с “разреженным” представлением объекта, хранимым в базе [18, 19].
4. Использование дискриминантной классификации [20, 21].

Для сравнения различных алгоритмов отслеживания объектов между собой было введено достаточно большое количество методов оценки и специфичных метрик оценки, которые учитывают и позволяют оценить наиболее частые варианты ошибок [13]:

- Отклонение: точность отслеживания ниже заявленной граничной точности.
- Ложноположительное срабатывание: отслеживается не тот объект.
- Ложноотрицательное срабатывание: нужный объект не отслеживается.

В существующих работах предложен ряд методов оценки, которые также можно разбить на два класса: с границей отсечения по точности и без такой границы [22]. Достаточно полный обзор существующих метрик можно найти в статье [13], тогда как в рамках данной работы необходимо отметить следующие недостатки существующих метрик:

- Все представленные на текущий момент метрики не учитывают производительность алгоритмов отслеживания. Сравнение скорости работы алгоритма проводится отдельно от оценки его точности, и зачастую эмпирически. Например, [23] рассматривает только количество кадров, на которых алгоритм отследил объекта, а в [24] дополнительно высчитывается дистанция между искомым и найденными объектами как мера точности.

- В работах не приводится сопоставление результатов работы с установленными параметрами алгоритмов. Таким образом, не происходит подбор параметров под наиболее оптимальные значения метрик, что не позволяет сделать вывод о применимости того или иного алгоритма или его модификации.

- В работах не учитывается, что определенный процент ошибочной работы алгоритмов отслеживания связан с результатами работы алгоритмов поиска. То есть ошибка отслеживания при использовании представленных метрик включает в себя ошибку поиска. Соответственно, представленная в [13] типизация ошибок не полна и требуется ее структурирование и переработка.

Таким образом, в рамках работы ставится задача разработки алгоритма и метрики для этого алгоритма, которая бы позволяла максимизировать как точность, так и производительность используемых алгоритмов отслеживания за счет настройки параметров конкретных алгоритмов отслеживания.

Для практической отработки предложенных далее в работе методов сравнения предлагается использовать простой в реализации алгоритм отслеживания в видеопотоке объектов по шаблону. Этот алгоритм является по сути адаптацией алгоритма поиска по шаблону на изображениях для отслеживания объектов в видеопотоке. Постановка задачи поиска объекта на изображении по шаблону [24] выглядит следующим образом.

Пусть дано изображение F , представляющее собой матрицу размером $w \times h$, где w – количество элементов в матрице (элемент матрицы соответствует пикселю на изображении) по горизонтали, h – количество элементов в матрице по вертикали (ширина и высота соответственно). Также предположим, что существует изображение T , которое является шаблоном искомого объекта, размер которого $l \times e$, $l \times w$, $e \times h$. Функцией F_{xy} обозначим пиксель, который является крайним левым пикселем подизображения F^T – изображения размера $l \times e$ на изображении F . При этом $0 \leq x \leq w - l$, $0 \leq y \leq h - e$. С помощью $R(I_1, I_2)$ обозначим функцию от двух изображений одинакового размера, которая позволяет рассчитать количественное значение разницы между этими изображениями. Возможные метрики для расчета количественного значения разности представлены в источниках [25–29]. Кроме того, в работах [26, 30] показано, что наиболее эффективная метрика – метрика нормированной корреляции:

$$R(I_1, I_2) = \frac{\sum_{x,y} [(I_1(x,y) - M(I_1(x,y)))(I_2(x,y) - M(I_2(x,y)))]}{\sqrt{\sum_{x,y} (I_1(x,y) - M(I_1(x,y)))^2 \sum_{x,y} (I_2(x,y) - M(I_2(x,y)))^2}} \quad (1)$$

При этом значения метрики (1) лежат в интервале $[0, 1]$, где 0 соответствует полностью различным изображениям, а 1 полностью идентичным.

Результатом работы алгоритма отслеживания по шаблону будет являться крайний левый пиксель подизображения F^T на изображении F , для которого при использовании данной метрики результат сравнения максимален:

$$F_{xy} = \arg \max_{x,y} R(T, F^T). \quad (2)$$

Для работы с использованием нормированных метрик вводят границу истинности a . Граница истинности представляет собой характеристику, которая указывает, насколько найденное представление объекта и искомое представление объекта должны соответствовать с точки зрения метрики. Таким образом, к представленной выше функции добавляется ограничение:

$$R(T, F^T) > a \quad (3)$$

В случае, если ограничение (3) выполняется для конкретного T и F^T , объект считается отслеживаемым успешно, в противном случае объект считается потерянным.

В рамках данной работы на практике рассматриваются параметры алгоритма отслеживания по шаблону и его модификаций. Используемые модификации представлены в [31]. Все ключевые параметры основного алгоритма и модификаций представлены в табл. 1. Ставится задача подбора таких a, b, c, k (параметры алгоритма и модификаций), чтобы эффективность алгоритма отслеживания, учитывающая как точность, так и производительность, была максимальна.

МЕТОД ОЦЕНКИ ТОЧНОСТИ И ПРОИЗВОДИТЕЛЬНОСТИ АЛГОРИТМОВ ОТСЛЕЖИВАНИЯ ОБЪЕКТОВ

Как было сказано ранее, в рамках рассмотренной литературы не приводится обоснование того, что приведенный набор ошибочных событий является полным. Проанализируем пространство ситуаций.

Для этого составим таблицу 2 по принципу сопоставления наличия и отсутствия объекта на конкретном кадре видеопотока и положительным или отрицательным результатом работы алгоритма отслеживания. Кроме того, для визуального анализа изложенных в таблице ситуаций был составлен рис. 1. В каждом случае на рисунке размещено три слоя: верхний слой – слой, показывающий наличие или отсутствие реального объекта в видеопотоке; средний слой – слой, показывающий результаты работы алгоритмов поиска и отслеживания; нижний слой – слой, обо-

значающий корректную или некорректную работу алгоритма отслеживания или поиска.

Сопоставление рассмотренных типов ситуаций с реальными объектами в кадре позволяет выявить необозначенный в стандартном анализе вариант ошибки в работе алгоритма отслеживания (рис. 1). В случае номер № 4 возникает ситуация пятого типа: появление дубликата, суть которой – ошибка в работе алгоритма отслеживания, связанная с преждевременной потерей отслеживаемого объекта с последующим нахождением потерянного объекта алгоритмом поиска. Но так как объект был потерян, то вновь найденный объект уже считается за другой объект, что по сути является даже более значительной ошибкой в работе алгоритма, чем преждевременная потеря объекта.

На основании представленной типизации ситуаций был составлен алгоритм (табл. 3) расчета ошибок различного типа, базирующийся на сравнении результатов работы алгоритма с эталонными результатами. В основу алгоритма легло следующее математическое описание типов ситуаций.

Обозначим с помощью $x_{t,il}, y_{t,il}, x_{b,il}, y_{b,il}$ координаты верхнего левого и нижнего правого углов соответственно области, которая в рамках алгоритма считается областью нахождения искомого объекта. Объект o с номером i на кадре l считается найденным верно, если позиция объекта $o_{il} = \{x_{t,il}, y_{t,il}, x_{b,il}, y_{b,il}\}$ совпадает с позицией одного из эталонных (нижний суффикс et) объектов $o_{et,il} = \{x'_{t,il}, y'_{t,il}, x'_{b,il}, y'_{b,il}\}$ с номером j на текущем изображении l с учетом коэффициентов наложения (e_1, e_2, e_3, e_4), задаваемых до начала работы алгоритма (по модулю меньше единицы). Таким образом, нужно ввести логическую функцию от координат искомого и эталонного объектов. В рамках данной работы предлагается взять следующую логическую функцию:

$$\begin{aligned} f(o_{il}, o_{et,jl}) &= f(x_{t,il}, y_{t,il}, x_{b,il}, y_{b,il}, x'_{t,jl}, y'_{t,jl}, x'_{b,jl}, y'_{b,jl}) = \\ &= \left(\frac{|x_{t,il} - x'_{t,jl}|}{\max(|x_{t,il} - x_{b,il}|, |x'_{t,jl} - x'_{b,jl}|)} < e_1 \right) \& \\ &\& \left(\frac{|y_{t,il} - y'_{t,jl}|}{\max(|y_{t,il} - y_{b,il}|, |y'_{t,jl} - y'_{b,jl}|)} < e_2 \right) \& \\ &\& \left(\frac{|x_{b,il} - x'_{b,jl}|}{\max(|x_{t,il} - x_{b,il}|, |x'_{t,jl} - x'_{b,jl}|)} < e_3 \right) \& \\ &\& \left(\frac{|y_{b,il} - y'_{b,jl}|}{\max(|y_{t,il} - y_{b,il}|, |y'_{t,jl} - y'_{b,jl}|)} < e_4 \right). \end{aligned} \quad (4)$$

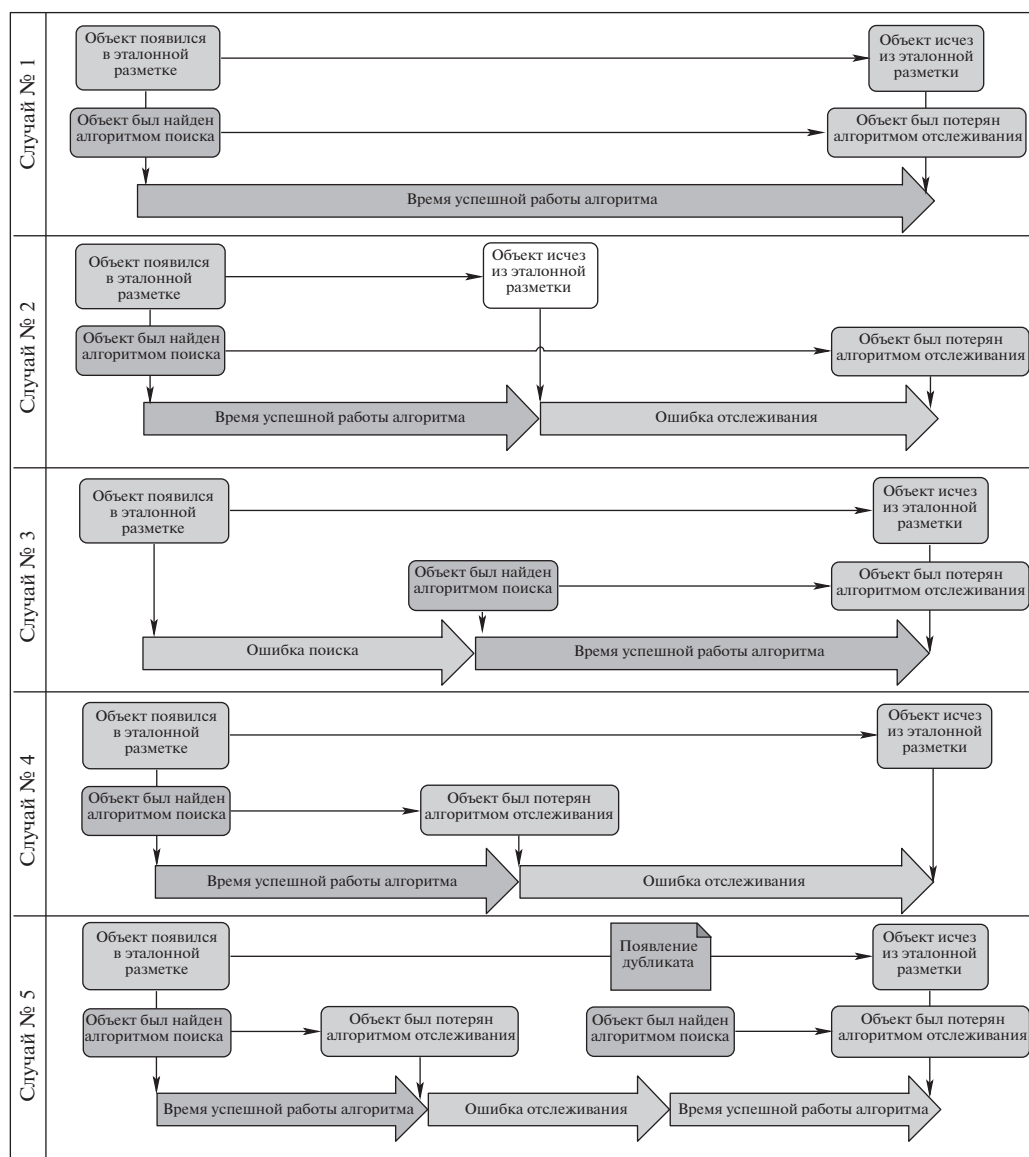


Рис. 1. Графическое сопоставление ситуаций работы алгоритма отслеживания и поведения модельных (реальных) объектов.

На основании предложенной классификации введем математическое описание ситуаций, которое впоследствии будет использовано при расчете статистики согласно пункту 9 представленного алгоритма в табл. 3.

• Тип № 1 – присутствующий объект o_{il} успешно отслеживается на выбранном кадре l , если объект успешно найден на рассматриваемом кадре или

$$\exists o_{i(l-1)}, o_{et,i(l-1)}, o_{et,il} : f(o_{i(l-1)}, o_{et,i(l-1)}) = 1 \ \& \ f(o_{il}, o_{et,il}) = 1. \quad (5)$$

• Тип № 2 – отслеживаемый объект o_{il} отсутствует на кадре l , если

$$\forall o_{et,jl} : f(o_{il}, o_{et,jl}) = 0. \quad (6)$$

• Тип № 3 – объект $o_{et,il}$ не отслеживается на кадре l , если

$$\forall o_{i(l-1)}, o_{et,i(l-1)}, o_{il} : f(o_{i(l-1)}, o_{et,i(l-1)}) = 0 \ \& \ f(o_{il}, o_{et,il}) = 0. \quad (7)$$

• Тип № 4 – $o_{et,il}$ не отслеживается на кадре l , если

$$\exists o_{i(l-1)}, o_{et,i(l-1)} : f(o_{i(l-1)}, o_{et,i(l-1)}) = 1 \ \& \ \nexists o_{jl} : f(o_{et,il}, o_{jl}) = 1. \quad (8)$$

Таблица 1. Ключевые параметры основного алгоритма и модификаций

№	Модификация	Обозначение параметра и его описание
1	Основной алгоритм	Граница истинности a в диапазоне $[0,1]$
2	Адаптация шаблона в процессе работы	Граница истинности a в диапазоне $[0,1]$ при использовании модификации Коэффициент сложения b имеющегося шаблона с изображением найденного объекта в диапазоне $[0,1]$
3	Повторный поиск потерянных объектов	Граница истинности a в диапазоне $[0,1]$ при использовании модификации
4	Дополнительная проверка объектов с помощью яркостной характеристики	Граница истинности a в диапазоне $[0,1]$ при использовании модификации Коэффициент сравнения гистограмм c имеющегося шаблона с изображением найденного объекта в диапазоне $[0,1]$
5	Уменьшение области поиска объекта	Коэффициент k расширения области поиска относительно размера искомого объекта

Таблица 2. Гипотетическое пространство ситуаций

	Объект присутствует в видеопотоке (существует)	Объект отсутствует в видеопотоке (не существует)
Объект отслеживается	Верная работа алгоритма отслеживания (тип № 1)	Неверная работа алгоритма отслеживания (тип № 2)
Объект не отслеживается	Объект не найден (тип № 3) Объект потерян (тип № 4)	Верная работа алгоритма отслеживания (не рассматривается)

Таблица 3. Алгоритм расчета результатов работы алгоритма отслеживания объектов

Алгоритм состоит из следующих элементов (по умолчанию ни у одного объекта, отслеживаемого в рамках программы нет ни отметки о потере, ни отметки об отслеживании):

- 1) Если есть доступные кадры в видеопотоке, то берется кадр, иначе переход к пункту 9.
- 2) Если для данного кадра есть сопоставленные, не рассмотренные ранее на этом кадре с ним эталонные объекты, то необходимо взять первый из них или следующий и перейти к пункту 3, иначе перейти к пункту 7.
- 3) Если для взятого эталона можно найти хотя бы один объект, удовлетворяющий критерию (4), то необходимо к эталону добавить отметку об отслеживании. Иначе перейти к пункту 5.
- 4) Если у эталона есть отметка о потере, то ее необходимо убрать и увеличить количество найденных дубликатов, иначе вернуться к пункту 2.
- 5) Если у эталонного объекта есть отметка об отслеживании, то ее необходимо убрать, добавить отметку о потере и увеличить время ошибки отслеживания для данного объекта, иначе перейти к пункту 6.
- 6) Необходимо увеличить время ошибки поиска для данного объекта и перейти к пункту 2.
- 7) Если для кадра остались не рассмотренные отслеживаемые объекты, то необходимо по очереди добавить для каждого из них время работы ошибки отслеживания. Иначе перейти к пункту 8.
- 8) Необходимо сохранить необходимые количественные данные и перейти к пункту 9.
- 9) Необходимо рассчитать статистику. Конец алгоритма.

• Тип № 5 – отслеживаемый объект o_{il} является дубликатом, если на одном из предыдущих кадров k

$$\exists o_{jk} : f(o_{jk}, o_{il}) = 1 \ \& \ f(o_{et, m(k-1)}, o_{j(k-1)}) = 1 \ \& \ f(o_{et, mk}, o_{jk}). \tag{9}$$

МЕТРИКА СРАВНЕНИЯ РЕЗУЛЬТАТОВ РАБОТЫ АЛГОРИТМОВ ОТСЛЕЖИВАНИЯ С УЧЕТОМ ТОЧНОСТИ И ПРОИЗВОДИТЕЛЬНОСТИ

Для оценки результатов работы алгоритмов отслеживания предлагается ввести следующие

статистические параметры, которые будут отражать результат работы алгоритма отслеживания:

- Общая ошибка в работе алгоритма поиска и отслеживания E_t ;

$L_{эт}$ — количество кадров, на которых находятся объекты в эталонной разметке;

L_T — количество кадров, на которых находятся объекты в результате работы алгоритмов поиска и отслеживания;

L_n — количество кадров, на которых находятся объекты в эталонной разметке и в результатах работы алгоритмов поиска и отслеживания.

$$E_t = \frac{E}{L_{э} + L_T - L_n} \times 100\% \quad (10)$$

- $\langle T \rangle$ — среднее время нахождения объекта в видеопотоке в рамках текущей итерации работы алгоритма отслеживания;

- $\langle T \rangle_{эт}$ — среднее время нахождения объекта в видеопотоке в рамках эталонной разметки текущего видеопотока;

$$\Delta T = |\langle T \rangle - \langle T \rangle_{эт}| \quad (11)$$

- D_n — количество дубликатов;

- $\max D$ — максимальное количество дубликатов объектов (случай № 5), которое было выбрано на основании перебора всех возможных параметров экспериментов;

- t_m — среднее время работы алгоритма на 1 кадр. Важно отметить, что берется именно среднее время работы на кадр, тогда как более традиционной является нормировка на один объект. Использование нормировки на объект не позволяет учесть так называемые “выпадающие” кадры — то есть кадры, которые не успевает обработать алгоритм отслеживания в связи с более высокой частотой их поступления, чем производительность алгоритма;

- N — выпадающие кадры;

- $\max t_m$ — максимальное время работы на 1 кадр в рамках всей серии экспериментов.

$$Q = \frac{\Delta T}{\langle T \rangle_{эт}} + E_t + \frac{D}{\max D} + \frac{t_m}{\max t_m} \quad (12)$$

Погрешность для значения критерия вычисляется следующим образом:

$$\Delta Q = Q \sqrt{\left(\frac{1}{fps \langle T \rangle} \right)^2 + \left(\frac{1}{fps \langle T_{ош} \rangle} \right)^2}, \quad (13)$$

где $\langle T \rangle_{ош}$ — среднее время ошибки работы алгоритмов отслеживания, fps — количество кадров в видеопотоке в секунду. Использование погрешности в расчете обусловлено тем, что предполагается использовать ручную разметку эталонного видеофайла.

Основной недостаток предлагаемого критерия — относительные значения, зависящие от выборки конкретного эксперимента. С его использованием нельзя провести абсолютное сравнение алгоритмов между собой.

МЕТОДИКА ПОСТРОЕНИЯ ЭКСПЕРИМЕНТА

Для проведения эксперимента было подготовлено специальное видео, устройство сцены которого и демонстрационные кадры из которого можно видеть на рис. 2. Продолжительность видео составила 1 минута 30 секунд, оно состоит из 2270 кадров. На видео присутствовало 41 лицо. Среднее время пребывания лица в кадре составило 3.3 секунды. Для оценки точности работы алгоритма была проведена ручная разметка видеоролика.

Для расчета указанных параметров был разработан программный комплекс, детальное описание работы которого находится за рамками данной работы.

Эксперименты в данной работе выстраиваются последовательно в соответствии с таблицей 1.

Граница истинности для алгоритма без модификаций

Перед проведением эксперимента в целях избегания потери кадров, связанных с большой зоной поиска объекта, был установлен коэффициент k для области поиска, равный трем. Результаты подбора параметра представлены в табл. 4. При этом видно, что с точки зрения критерия подбор параметров необходим для минимально эффективной работы алгоритма. В частности:

- Граница истинности в интервале от 0 до 0.8 приводит к потере обработанных кадров даже с учетом уменьшенной зоны поиска, что связано с тем, что отслеживаемые объекты “не теряются”, то есть идет отслеживание ложных объектов (ситуация типа 2). Устанавливать значения границы истинности из указанного диапазона не целесообразно в принципе. При этом критерий позволяет избежать установки таких параметров без подсчета количества потерянных кадров, что показывает его эффективность.

- Оптимальные значения границы истинности лежат в интервале от 0.9 до 0.94. Результаты в этих интервалах совпадают в пределах погрешности.

- Максимальная разница в относительной ошибке результатов работы алгоритма отслеживания составляет 1.92 раза.

- Абсолютные значения критерия различаются на порядки, как и абсолютная разница в работе алгоритма с разными параметрами, что показы-



Рис. 2а. Устройство сцены тестового видеоролика. 2б. Примеры кадров из видеоролика.

вает необходимость работы с параметрами алгоритма перед их практическим использованием.

Подбор оптимальных параметров для модификаций

На основании табл. 1 приводим в табл. 5 результаты подбора параметров для предложенных модификаций. Модификации добавляются по-

степенно в соответствии со строчками в таблице в формате “нарастающим итогом”. В табл. 5 в графах “ α ” и “доп. параметр” приводятся подобранные значения параметров, а в графах “ E_t ”, “ D ”, и “влияние на производительность” дополнительно указывается диапазон влияния изменения параметров на соответствующий показатель.

Таблица 4. Подбор параметров для алгоритма без модификаций

α	ΔT , мс	E_t	D	t , мс	Q	Δ	N
0.1	49691	72%	18	81.23	17.7	0.03	483
0.2	50977	71.70%	18	77.57	18.18	0.03	385
0.3	49362	72.20%	21	61.21	17.48	0.03	282
0.4	45502	76.60%	20	59.34	16.39	0.03	278
0.5	45574	75.50%	18	50.75	16.291	0.03	242
0.6	32414	67.00%	16	35.47	11.66	0.03	120
0.7	21897	68.70%	18	21.7	8.25	0.03	31
0.8	8873	57.70%	16	14.23	3.81	0.03	15
0.9	512	33.70%	16	11.1	0.91	0.03	0
0.91	297	39.00%	17	10.88	0.90	0.06	0
0.92	22	43.10%	19	10.73	0.89	0.05	0
0.93	438	39.90%	21	10.6	1.01	0.07	0
0.94	570	37.60%	18	10.51	0.97	0.11	0
0.95	1389	31.20%	25	10.51	1.27	0.14	0
0.96	1963	34.40%	17	10.51	1.34	0.13	0
0.97	1996	37.40%	29	10.51	1.58	0.15	0
0.98	2427	40.90%	51	10.51	2.10	0.21	0
0.99	2191	44.80%	26	10.51	1.66	0.15	0
1	Алгоритм не работает, объекты сразу же не отслеживаются.						

Таблица 5. Влияние параметров модификаций на производительность

Модификация	α	Доп. параметр	E_t	D	Масштабы влияния на производительность
Адаптация шаблона в процессе работы	0.955	0.3 – 0.6	21.90% 4-х влияние	11 4-х влияние	8-х влияние
Повторный поиск потерянных объектов	0.95	—	22.10% Отсутствует	6 2-х влияние	Увеличивает производительность в 2.4 раза
Проверка объектов с помощью яркостной характеристики	0.925	0.65–0.8	10.60% 10-х влияние	3 3-х влияние	Уменьшает производительность на 8%
Уменьшение области поиска объекта	0.925	2–3	9.00%	3	10-х влияние

Также в рамках работы целесообразно показать, что неверный подбор параметров может привести с точки зрения оценки ошибки работы алгоритма к отказу от использования эффективных модификаций, пример чего можно видеть на рис. 3 (ось Y – процент ошибки, ось X – граница истинности). В качестве примера модификации используется модификация адаптации шаблона. Как видно, при значении границы истинности от 0 до 0.92 использование модификации нецелесообразно, тогда как при значении параметра от 0.93 до 1 использование модификации целесообразно и дает меньший процент ошибок, чем немодифицированный вариант.

ЗАКЛЮЧЕНИЕ

В работе наглядно показано, что использование параметрической оптимизации с целью повышения эффективности работы алгоритмов отслеживания объектов является не просто полезным для повышения точности работы алгоритмов, но на самом деле необходимым для обеспечения работы рассматриваемых алгоритмов.

Основные выводы по итогам работы:

- Метрики для оценки качества работы алгоритмов отслеживания должны включать такой параметр, как производительность алгоритма на один кадр. Отсутствие этого параметра может приводить к неучету алгоритмом ряда кадров, которые не успели обработать в рамках алгоритма. Потеря кадров ведет к дополнительной ошибке, которая не зависит от исходного алгоритма отслеживания объектов.

- Неверный подбор параметров может привести, с точки зрения оценки ошибки работы алгоритма, к отказу от использования эффективных модификаций.

- Для алгоритма отслеживания по шаблону при использовании нормализованной корреляции оптимально использовать границу отсечения в интервале от 0.92 до 0.955 во всех рассматриваемых случаях.

- С помощью подбора параметров можно уменьшить процент ошибок в несколько раз. В данной статье показано, что для алгоритма отслеживания объектов по шаблону можно увеличить точность в 4.7 раза при увеличении производительности в 2.2 раза при использовании модификаций и подборе их параметров.

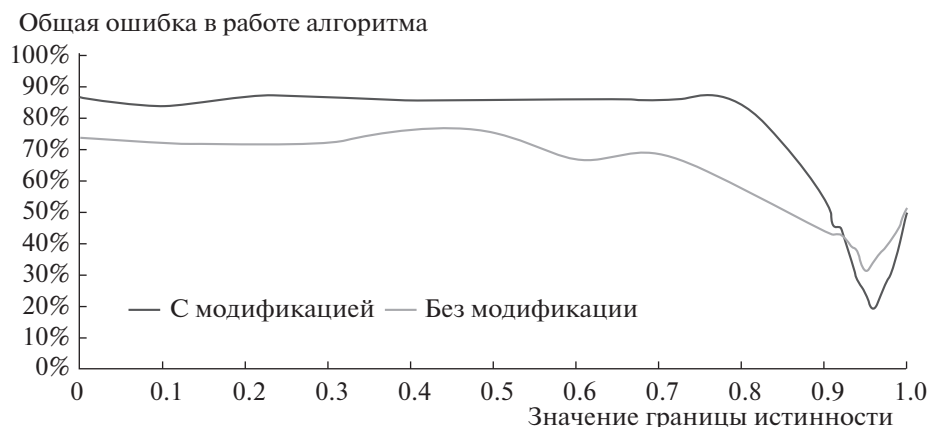


Рис. 3. Сравнение ошибки работы алгоритмов отслеживания в зависимости от границы истинности.

Дальнейшее направление работы заключается в исследовании существующих алгоритмов отслеживания объектов и их модификаций с целью выявления эффективных и неэффективных, а также с целью подбора параметров для существующих алгоритмов отслеживания объектов. Направления исследования: перебор различных алгоритмов отслеживания; создание и использование видеороликов, снятых в различных условиях внешней среды, с целью подбора параметров под среду; разработка модификаций для алгоритмов отслеживания и их типизация на основании представленной в работе классификации; анализ влияния количества объектов на результаты работы алгоритма и т.п.

БЛАГОДАРНОСТИ

Исследование выполнено в рамках проекта № 25.2911.2017/4.6 “Апробация технологии снижения риска возникновения и уменьшения последствий катастроф техногенного происхождения за счет минимизации влияния человеческого фактора на надежность и безаварийность работы АЭС и других опасных объектов” в соответствии с конкурсной частью государственного задания НИЯУ МИФИ на 2017–2019 гг.

СПИСОК ЛИТЕРАТУРЫ

1. *Bradski G., Kaehler A.* Learning OpenCV: Computer vision with the OpenCV library. “O’Reilly Media, Inc.”, 2008.
2. *Al-Kaff A. et al.* Survey of computer vision algorithms and applications for unmanned aerial vehicles // *Expert Systems with Applications*. 2018. Т. 92. С. 447–463.
3. *Buch N., Velastin S.A., Orwell J.* A review of computer vision techniques for the analysis of urban traffic // *IEEE Transactions on Intelligent Transportation Systems*. 2011. Т. 12. № 3. С. 920–939.
4. *Alyushin M.V., Alyushin V.M., Kolobashkina L.V.* “Methodological aspects of automated forecasting of emergency situations of technogenic origin”. *Questions of psychology*. 2016. V. 2. P. 83–90.
5. *Alyushin M.V., Alyushin A.V., Belopolskiy V.M., Kolobashkina L.V., Ushakov V.D.* “Optical Technologies for Monitoring Systems of the Current Functional State of the Operational Management Personnel of Nuclear Power Facilities”. *Global Nuclear Safety*. 2013. V. 2. № 7. P. 69–77.
6. *Joskowicz L.* Future Perspectives on Statistical Shape Models in Computer-Aided Orthopedic Surgery: Beyond Statistical Shape Models and on to Big Data // *Computer Assisted Orthopaedic Surgery for Hip and Knee*. Springer, Singapore, 2018. С. 199–206.
7. *Куций Н.Н., Осипова Е.А.* Параметрическая оптимизация систем с интегральной широтно-импульсной модуляцией с использованием беспoiseкового градиентного алгоритма // *Приборы и системы. Управление, контроль, диагностика*. 2012. № 1. С. 2–6.
8. *Кохно А.Г.* Многокритериальная параметрическая оптимизация судовых автоматизированных систем: дисс. канд. техн. наук // Санкт-Петербургский государственный университет водных коммуникаций, Санкт-Петербург. 2012.
9. *Скляренко А.А.* Методы решения задачи пописельной s-аппроксимации мультитоновых изображений и их оптимизация // Ростов-на-Дону: Издательский центр ДГТУ. 2012.
10. *Dunai Dunai L. et al.* Euro banknote recognition system for blind people // *Sensors*. 2017. Т. 17. № 1. С. 184.
11. *Bruce B.R., Aitken J.M., Petke J.* Deep parameter optimization for face detection using the viola-jones algorithm in OpenCV // *International Symposium on Search Based Software Engineering*. Springer, Cham. 2016. С. 238–243.
12. *Egorov A.D., Shtanko A.N., Minin P.E.* Selection of Viola–Jones algorithm parameters for specific conditions // *Bulletin of the Lebedev Physics Institute*. 2015. Т. 42. № 8. С. 244–248.
13. *Smeulders A.W.M. et al.* Visual tracking: An experimental survey // *IEEE transaction on pattern analysis and machine intelligence*. 2014. Т. 36. № 7. С. 1442–1468.
14. *Oron S. et al.* Locally orderless tracking // *International Journal of Computer Vision*. 2015. Т. 111. № 2. С. 213–228.
15. *Briechele K., Hanebeck U.D.* Template matching using fast normalized cross correlation // *Optical Pattern Recognition XII. International Society for Optics and Photonics*, 2001. Т. 4387. С. 95–103.
16. *Ross D.A. et al.* Incremental learning for robust visual tracking // *International journal of computer vision*. 2008. Т. 77. № 1–3. С. 125–141.
17. *Isard M., Blake A.* A mixed-state condensation tracker with automatic model-switching // *Computer Vision, 1998. Sixth International Conference on. IEEE*, 1998. С. 107–112.
18. *Mei X., Ling H.* Robust visual tracking using ℓ_1 minimization // *Computer Vision, 2009 IEEE 12th International Conference on. IEEE*, 2009. С. 1436–1443.
19. *Ayvaci A., Raptis M., Soatto S.* Sparse occlusion detection with optical flow // *International journal of computer vision*. 2012. Т. 97. № 3. С. 322–338.
20. *Nguyen H.T., Smeulders A.W.M.* Robust tracking using foreground-background texture discrimination // *International Journal of Computer Vision*. 2006. Т. 69. № 3. С. 277–293.
21. *Godec M., Roth P.M., Bischof H.* Hough-based tracking of non-rigid objects // *Computer Vision and Image Understanding*. 2013. Т. 117. № 10. С. 1245–1256.
22. *Salti S., Cavallaro A., Di Stefano L.* Adaptive appearance modeling for video tracking: Survey and evaluation // *IEEE Transactions on Image Processing*. 2012. Т. 21. № 10. С. 4334–4348.
23. *Chan T.F., Vese L.A.* Active contours without edges // *IEEE Transactions on image processing*. 2001. Т. 10. № 2. С. 266–277.
24. *Briechele K., Hanebeck U.D.* Template matching using fast normalized cross correlation // *Optical Pattern Recognition XII. International Society for Optics and Photonics*, 2001. Т. 4387. С. 95–103.

25. Shen X. et al. Detecting and aligning faces by image retrieval // Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on. IEEE, 2013. С. 3460–3467.
26. Leibe B., Leonardis A., Schiele B. Robust object detection with interleaved categorization and segmentation // International journal of computer vision. 2008. Т. 77. № 1–3. С. 259–289.
27. Ballard D.H. Generalizing the Hough transform to detect arbitrary shapes // Readings in computer vision. 1987. С. 714–725.
28. Felzenszwalb P.F. et al. Object detection with discriminatively trained part-based models // IEEE transactions on pattern analysis and machine intelligence. 2010. Т. 32. № 9. С. 1627–1645.
29. Chen D. et al. Joint cascade face detection and alignment // European Conference on Computer Vision. Springer, Cham, 2014. С. 109–122.
30. Martins P., Caseiro R., Batista J. Generative face alignment through 2.5 d active appearance models // Computer Vision and Image Understanding. 2013. Т. 117. № 3. С. 250–268.
31. Samoylov A.S. et al. Face tracking for a system of collecting statistics on visitors and quality assessment of its functioning // Journal of Theoretical & Applied Information Technology. 2015. Т. 71. № 3.

Vestnik Natsional'nogo issledovatel'skogo yadernogo universiteta "MIFI", 2019, vol. 8, no. 3, pp. 352–362

Optimization of Parametric Object Video Stream Tracking Algorithms

A. D. Egorov[#]

National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Moscow, 115409 Russia

[#]e-mail: ADEgorov@mephi.ru

Received January 15, 2019; revised February 28, 2019; accepted March 12, 2019

Abstract—An increase in the efficiency of algorithms of object tracking in video streams through the parametric optimization of these algorithms has been discussed. In contrast to most of such studies, accuracy and performance parameters are taken into account in one objective optimization function. There are five different results of object tracking in video streams: correct tracking, false object tracking, premature object loss, finding an object with a delay, and double object detection after premature loss. It has been shown that the parametric optimization makes it possible to increase the accuracy and performance of template matching tracking by factors of 4.7 and 2.2, respectively. At the same time, it has been found that the absence of parametric optimization can lead in practice to the refusal of efficient modifications. For example, template adaptation modification increases the template tracking quality only in 10% of the cases and decreases it in other 90% of the cases. Particular parameters of the algorithm for template object tracking in video streams have been recommended for computer vision systems.

Keywords: parametric optimization, object tracking in video stream, template matching, computer vision, video stream

DOI: 10.1134/S2304487X19030052

REFERENCES

1. Bradski G., Kaehler A. Learning OpenCV: Computer vision with the OpenCV library. "O'Reilly Media, Inc.", 2008.
2. Al-Kaff A. et al. Survey of computer vision algorithms and applications for unmanned aerial vehicles. *Expert Systems with Applications*, 2018, pp. 447–463.
3. Buch N., Velastin S.A., Orwell J. A review of computer vision techniques for the analysis of urban traffic. *IEEE Transactions on Intelligent Transportation Systems*, 2011, vol. 12, no. 3, pp. 920–939.
4. Alyushin M.V., Alyushin V.M., Kolobashkina L.V. Methodological aspects of automated forecasting of emergency situations of technogenic origin. *Questions of psychology*, 2016, vol. 2, pp. 83–90.
5. Alyushin M.V., Alyushin A.V., Belopolskiy V.M., Kolobashkina L.V., Ushakov V.D. Optical Technologies for Monitoring Systems of the Current Functional State of the Operational Management Personnel of Nuclear Power Facilities. *Global Nuclear Safety*, 2013, vol. 2(7), pp. 69–77.
6. Joskowicz L. Future Perspectives on Statistical Shape Models in Computer-Aided Orthopedic Surgery: Beyond Statistical Shape Models and on to Big Data. *Computer Assisted Orthopaedic Surgery for Hip and Knee*. Springer, Singapore, 2018, pp. 199–206.
7. Kutsiy N.N. Parametricheskaya optimizatsiya sistem s integral'noj shirotno-impul'snoy modulyaciej s ispol'zovaniem bespoiskovogo gradientnogo algoritma [Parametric optimization of systems with integral pulse-width modulation using a searchless gradient algorithm].

- Instruments and Systems: Monitoring, Control, and Diagnostics, 2012, vol. 1, pp. 2–6.
8. Kohno A.G. Mnogokriterial'naya parametriceskaya optimizaciya sudovyh avtomatizirovannyh sistem. Kand. Diss. [*Multi-criteria parametric optimization of ship automated systems*. Kand. Diss.]. Saint-Petersburg, 2012.
 9. Sklyarko A.A. Metody resheniya zadachi popiksel'noj s-approksimacii mul'titonovyh izobrazhenij i ih optimizaciya. Kand. Diss. [*methods for solving the pixix s-approximation multiton images and their optimization*. Kand. Diss.], Rostov-on-Don, 2012.
 10. Dunai Dunai L. et al. Euro banknote recognition system for blind people. *Sensors*, 2017, vol. 17(1), pp. 184.
 11. Bruce B.R., Aitken J.M., Petke J. Deep parameter optimization for face detection using the viola-jones algorithm in OpenCV. *International Symposium on Search Based Software Engineering*. Springer, Cham, 2016, pp. 238–243.
 12. Egorov A.D., Shtanko A.N., Minin P.E. Selection of Viola–Jones algorithm parameters for specific conditions. *Bulletin of the Lebedev Physics Institute*, 2015, vol. 42(8), pp. 244–248.
 13. Smeulders A.W.M. et al. Visual tracking: An experimental survey. *IEEE transactions on pattern analysis and machine intelligence*, 2014, vol. 36(7), pp. 1442–1468.
 14. Oron S. et al. Locally orderless tracking. *International Journal of Computer Vision*, 2015, vol. 111(2), pp. 213–228.
 15. Briechele K., Hanebeck U.D. Template matching using fast normalized cross correlation. *Optical Pattern Recognition XII*. International Society for Optics and Photonics, 2001, vol. 4387, pp. 95–103.
 16. Ross D.A. et al. Incremental learning for robust visual tracking. *International journal of computer vision*, 2008, vol. 77 (1–3), pp. 125–141.
 17. Isard M., Blake A. A mixed-state condensation tracker with automatic model-switching. *Computer Vision, 1998. Sixth International Conference on. IEEE*, 1998, pp. 107–112.
 18. Mei X., Ling H. Robust visual tracking using ℓ_1 minimization. *Computer Vision, 2009, IEEE 12th International Conference on. IEEE*, 2009, pp. 1436–1443.
 19. Ayvaci A., Raptis M., Soatto S. Sparse occlusion detection with optical flow. *International journal of computer vision*. 2012, vol. 97(3), pp. 322–338.
 20. Nguyen H.T., Smeulders A.W.M. Robust tracking using foreground-background texture discrimination. *International Journal of Computer Vision*, 2006, vol. 69(3), pp. 277–293.
 21. Godec M., Roth P.M., Bischof H. Hough-based tracking of non-rigid objects. *Computer Vision and Image Understanding*, 2013, vol. 117(10), pp. 1245–1256.
 22. Salti S., Cavallaro A., Di Stefano L. Adaptive appearance modeling for video tracking: Survey and evaluation. *IEEE Transactions on Image Processing*, 2012, vol. 21(10), pp. 4334–4348.
 23. Chan T.F., Vese L.A. Active contours without edges. *IEEE Transactions on image processing*, 2001, vol. 10(2), pp. 266–277.
 24. Briechele K., Hanebeck U.D. Template matching using fast normalized cross correlation. *Optical Pattern Recognition XII*, 2001, vol. (4387), pp. 95–103.
 25. Shen X. et al. Detecting and aligning faces by image retrieval. *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, 2013, pp. 3460–3467.
 26. Leibe B., Leonardis A., Schiele B. Robust object detection with interleaved categorization and segmentation // *International journal of computer vision*, 2008, vol. 77(1–3), pp. 259–289.
 27. Ballard D.H. Generalizing the Hough transform to detect arbitrary shapes. *Readings in computer vision*, 1987, pp. 714–725.
 28. Felzenszwalb P.-F. et al. Object detection with discriminatively trained part-based models. *IEEE transactions on pattern analysis and machine intelligence*, 2010, vol. 32(9), pp. 1627–1645.
 29. Chen D. et al. Joint cascade face detection and alignment. *European Conference on Computer Vision*, 2014, pp. 109–122.
 30. Martins P., Caseiro R., Batista J. Generative face alignment through 2.5 d active appearance models. *Computer Vision and Image Understanding*, 2013, vol. 117(3), pp. 250–268.
 31. Samoylov A.S. et al. Face tracking for a system of collecting statistics on visitors and quality assessment of its functioning. *Journal of Theoretical & Applied Information Technology*, 2015, vol. 71(3).