

Слежение за объектами с помощью адаптивных обобщенных корреляционных фильтров

В.И. Кобер*, А.Н. Ручай**, В.Н. Карнаухов*

* *Институт проблем передачи информации, Российская академия наук, Москва, Россия*

** *Челябинский государственный университет, Челябинск, Россия*

Поступила в редколлегию 15.04.2019

Аннотация—Предложен алгоритм слежения за несколькими объектами, использующий локально-адаптивную обобщенную корреляционную фильтрацию. Алгоритм слежения является устойчивым к геометрическим преобразованиям объектов, частичному перекрытию объектов предметами сцены, неравномерному освещению сцены и наличию аддитивного шума на изображениях сцены. Предложенная система использует обобщенные оптимальные корреляционные фильтры и схему предсказания, построенную на основе кинематической модели движения объектов. Адаптация корреляционных фильтров к текущим изменениям сцены производится с помощью их итеративного обучения. Работа предложенного алгоритма сравнивается с работой известных алгоритмов видео слежения на известных базах данных с точки зрения общепринятых метрик качества слежения и скорости обработки.

КЛЮЧЕВЫЕ СЛОВА: слежение за объектами, адаптивный фильтр, корреляционный фильтр.

1. ВВЕДЕНИЕ

В настоящее время слежение за объектами является популярной темой в инженерном и компьютерном зрении [1, 2]. Примерами известных приложений слежения за объектами являются видеонаблюдение, автомобильная навигация, взаимодействие человека с компьютером и робототехника [3].

Основной сложностью задачи слежения за объектами является то, что наблюдаемая сцена может динамически искажаться: быть искажена аддитивным шумом, иметь сложный, меняющийся фон и неравномерное освещение; кроме того, объекты слежения могут быть подвержены геометрическим изменениям, таким как изменение формы, масштабирование, частичное перекрытие другими предметами сцены и т.д. [4]. Требование работы системы в реальном времени дополнительно усложняет задачу разработки эффективных алгоритмов слежения.

В этой статье мы определяем задачу слежения как обнаружение объектов и вычисление их положения в каждом кадре видео последовательности. Объекты для слежения задаются в первом кадре с помощью ограничивающих прямоугольных рамок. В последнее время было предложено много интересных алгоритмов с целью решения данной задачи [5].

Для решения проблемы частичного перекрытия объектов и геометрических искажений сцены был предложен алгоритм слежения, построенный на взаимодействии глобального фильтра и нескольких локальных фильтров, учитывающих возможную деформацию объектов [6]. Динамическое обновление модели объектов, применение адаптивного порога и меняющегося размера ядра в быстром алгоритме среднего смещения были рассмотрены в работе [7]. Кроме того, авторы [8] предложили аналитическую модель неизменяемых частей объектов на основе нового корреляционного фильтра с ядром (KCF) и фильтра двойной корреляции (DCF) для многоканального обобщения линейных корреляционных фильтров.

Для решения проблемы изменения масштаба был предложен метод, использующий сверточные сети и корреляционный фильтр [9, 10]. В работах [11, 12] был предложен алгоритм на основе RFC фильтрации, в котором строится пространственно-временная матрица углов для выбора надежных частей (фрагментов) объектов с помощью нескольких корреляционных фильтров и последующего поиска оптимального кандидата для вычисления аффинного преобразования объекта с наибольшей вероятностью. В статье [13] были предложены масштабно-пространственный фильтр и многоблочная схема на основе KCF с динамической моделью, которая аппроксимирует изменения состояния перекрытия и деформации.

Для решения проблемы размытия изображения сцены вследствие быстрого движения объектов был предложен метод регуляризации [14]. Для решения проблемы быстрого исчезновения объектов из кадра в работе [15] объединили технику пространственного управления окнами с древовидной структурой, которая эффективно регистрирует появления объектов в узлах дерева.

В статье [16] представлена система слежения, использующая корреляционные фильтры и надежное предсказание Калмана для определения положения объекта. В работе использовался адаптивный подход к улучшению коэффициента дискриминации между объектами и подавлению сложного фона. Основным недостатком этого алгоритма является отсутствие устойчивости к геометрическим изменениям формы объектов, частичному перекрытию объектов и неравномерному освещению сцены.

Методы распознавания на основе сопоставления с шаблоном без обучения неэффективны при искажении функций контура и формы объектов [17, 18].

В работе [19] предложена система слежения, использующая банк композитных корреляционных фильтров со сложными ограничениями. Система разделяет текущее изображение сцены на несколько фрагментов с использованием модели кинематического предсказания положения объекта. Фрагменты обрабатываются независимо с помощью адаптивных фильтров. Фильтры синтезируются итеративным алгоритмом, который оптимизирует коэффициент дискриминации для каждого объекта слежения. Несколько объектов на исходной сцене могут быть обнаружены и классифицированы путем анализа распределений интенсивности и фазы на выходной комплексной корреляционной плоскости для каждого фрагмента.

В статье [20] слежение за объектом выполняется с помощью детектора особых точек и банка локально-адаптивных корреляционных фильтров. Особые точки позволяют выделять локальные области на наблюдаемой сцене вокруг потенциального положения объекта. Банк корреляционных фильтров позволяет надежно обнаруживать объекты и вычислять оценки положения объектов в пределах сцены путем обработки выделенных локальных областей.

Авторы [21, 22] предложили алгоритм слежения за лицами, основанный на локально-адаптивной корреляционной фильтрации. Этот алгоритм способен следить за лицом при небольших изменениях позы, формы, масштаба и частичном перекрытии. Система распознавания адаптируется в режиме онлайн с использованием информации о текущих и прошлых кадрах сцены. Адаптивный композитный фильтр инвариантный к геометрическим изменениям синтезируется путем объединения нескольких оптимальных фильтров, предназначенных для распознавания различных вариантов изображения лица. Основным недостатком этого алгоритма является неустойчивость к фоновым изменениям.

В статье [23] представлен адаптивный подход к синтезу фильтров для слежения за объектами. Этот алгоритм является адаптивным к изменению формы объектов, неравномерному освещению сцены, аддитивному шуму и частичному перекрытию. Однако процедура повторной инициализации, которая необходима при быстром изменении формы объекта, является довольно трудоемкой. Кроме того, алгоритм учитывает изменения только вокруг объекта, а

не на всем фоне. Отметим, что используемое рекурсивное вычисление адаптивного фильтра не является надежным при динамическом распознавании объектов.

Авторы в работе [24] предложили алгоритм распознавания частично закрытых объектов, состоящий из набора композиционных адаптивных корреляционных фильтров. Алгоритм является адаптивным к входной сцене, которая может содержать объекты, фрагменты объектов, ложные объекты и сложный фон. На основе решения оптимизационной задачи в процессе обучения в статье [25] был предложен эффективный алгоритм распознавания частично перекрытых объектов.

В этой работе мы предлагаем алгоритм слежения за объектами на основе банка локально-адаптивной оптимальной корреляционной фильтрации. Алгоритм начинается с процедуры инициализации, в которой интерактивно выбираются объекты для слежения. Для каждого выбранного объекта синтезируется обобщенный оптимальный фильтр для надежного обнаружения объекта в видеокадре. Стандартные корреляционные фильтры синтезируются при условии, что объект распознавания точно известен. Однако в реальной жизни объект распознавания часто задан зашумленным изображением, содержащим эталон в произвольном месте. В этом случае для синтеза традиционного фильтра используют интерактивные методы сегментирования и выделения эталона. В этой статье используется модель сигнала с неявным заданием объекта на сложном фоне, а оценки функций объектов и их формы получаются автоматически из зашумленных изображений объектов. Таким образом, из процесса распознавания исключается этап ручной подготовки объектов посредством сегментирования. Предложенный алгоритм включает в себя процедуру автоматической повторной инициализации, которая восстанавливает слежение за нужными объектами в случае их потери. Импульсные характеристики обобщенных оптимальных фильтров используются для синтеза композитных фильтров, инвариантных к геометрическим искажениям объектов.

Статья организована следующим образом. Раздел 2 описывает процедуру синтеза композитного линейного фильтра на основе обобщенных оптимальных корреляционных фильтров, рассчитанных для неперекрывающейся модели сигнала. Раздел 3 описывает предложенный алгоритм слежения за объектами с помощью локально-адаптивной корреляционной фильтрации. Результаты компьютерного моделирования представлены в разделе 4. Наконец, в разделе 5 даны наши выводы.

2. СИНТЕЗ КОМПОЗИТНОГО ЛИНЕЙНОГО ФИЛЬТРА

Объекты выбираются в качестве цели слежения на изображении первого кадра видео последовательности. Предположим, что объект распознавания задан зашумленным изображением, содержащим эталон в произвольном месте. Для того чтобы синтезировать корреляционный фильтр, необходимо выделить объект из изображения эталона, что позволит получить точную функцию яркости и форму объекта. В работе [26, 27] был предложен синтез обобщенного оптимального фильтра для обнаружения объекта, который неявно задан зашумленным изображением.

Рассмотрим математическую модель и синтез оптимального фильтра для объекта, который врезан в случайный фон. Одномерное обозначение используется для простоты. Обозначим изображение эталона как $r(x)$, бинарную функцию обратной поддержки объекта как $\mathbf{w}(x)$. Функция принимает значение единицы вне области объекта и ноль внутри области. Пусть $s(x)$ — наблюдаемая сцена, представленная моделью неперекрывающегося сигнала, то есть объект врезан в фоновое изображение. Изображение сцены содержит также аддитивный шум датчиков. Модели изображения эталона и наблюдаемой сцены формально определяется как

$$r(x) = t(x - x_r) + b_r(x)\mathbf{w}(x - x_r) + n_r(x), \quad (1)$$

$$s(x) = t(x - x_s) + b_s(x)\mathbf{w}(x - x_s) + n_s(x), \quad (2)$$

где $b_r(x)$ и $b_s(x)$ — функции фона на изображении эталона и наблюдаемой сцене, соответственно; $n_r(x)$ и $n_s(x)$ — аддитивный шум на изображении эталона и наблюдаемой сцене, соответственно; $N_r(\omega)$ и $N_s(\omega)$ — спектральные плотности мощности шума на изображении эталона и наблюдаемой сцене, соответственно; x_r и x_s — координаты объекта на изображении эталона и наблюдаемой сцене, соответственно; $b_r(x)$ и $b_s(x)$ — стационарные случайные процессы с математическими ожиданиями μ_r и μ_s , описывающие фон на изображении эталона и наблюдаемой сцене, соответственно; $B_r^0(\omega)$ и $B_s^0(\omega)$ — их спектральные плотности мощности, соответственно. Все случайные процессы и случайные величины являются статистически независимыми друг от друга.

Если объект и его бинарная функция поддержки явно заданы, то частотная характеристика обобщенного оптимального фильтра (GOF) является наилучшей с точки зрения отношения пика корреляции к энергии корреляционной плоскости (РОЕ) [26]:

$$H(\omega) = \frac{[T(\omega) + \mu_r W(\omega)]R^*(\omega)}{|T(\omega) + \mu_r W(\omega)|^2 + (2\pi)^{-1}B_r^0(\omega) \bullet |W(\omega)|^2 + N_r(\omega)} \times \frac{T^*(\omega) + \mu_s W^*(\omega)}{|T(\omega) + \mu_s W(\omega)|^2 + (2\pi)^{-1}B_s^0(\omega) \bullet |W(\omega)|^2 + N_s(\omega)} \quad (3)$$

где $T(\omega)$, $W(\omega)$ и $R(\omega)$ — преобразования Фурье от $t(x)$, $\mathbf{w}(x)$ и $r(x)$, соответственно; \bullet обозначает операцию свертки, $*$ — комплексное сопряжение.

Предположим, что объект $t(x, y)$ находится внутри фрагмента изображения $f_r(x, y)$. Ковариационная функция фрагмента, искаженного белым аддитивным шумом, может быть вычислена как $C(x, y) = C_{f_r}(x, y) + C_n(x, y)$, где $C_n(x, y) = \sigma_n^2 \delta(x, y)$ — ковариационная функция аддитивного шума, σ_n^2 — дисперсия белого шума, $\delta(x, y)$ — дельта функция Кронекера, $C_{f_r}(x, y)$ — ковариационная функция неискаженного сигнала. Дисперсия шума может быть оценена как

$$\sigma_n^2 = C(0, 0) - C_{f_r}(0, 0),$$

где значение $C_{f_r}(0, 0)$ может быть приближенно вычислено простой линейной экстраполяцией как

$$C_{f_r}(0, 0) = 2C(0, 1) - C(0, 2).$$

Полученная оценка дисперсии используется для оценки спектральной плотности мощности белого шума на изображениях эталона и сцены. В формуле (1) объект задается неявно с помощью изображения эталона. Для реализации фильтра необходимы преобразования Фурье объекта распознавания и функции поддержки объекта, что противоречит постановке задачи, что информация об объекте недоступна. Для оценивания необходимых параметров будем использовать только изображение эталона. Хорошая оценка [26] обратной функции поддержки объекта может быть получена с помощью пороговой обработки с минимальной ошибкой классификации изображения эталона после Винеровской фильтрации. Два способа оценивания функции объекта распознавания в зависимости от статистик первого и второго порядка на изображении эталона после Винеровской фильтрации были предложены в работе [26]. Подставляя эти оценки в (3), получим приближенные варианты оптимального фильтра GOF, которые затем будем использовать для синтеза композитных адаптивных фильтров.

Пусть $h_i(x)$ — импульсная характеристика фильтра GOF для i -ой версии объекта распознавания $t_i(x)$ размером d элементов. Пусть $H = \{h_i(x); i = 1, \dots, N\}$ — множество импульсных

характеристик GOF, рассчитанных для обучающих версий объекта $T = \{t_i(x); i = 1, \dots, N\}$. Кроме того, пусть $S = \{s_i(x); i = 1, \dots, M\}$ — набор ложных объектов. Линейный композитный фильтр $p(x)$, который способен распознать все версии объекта из T с помощью оптимальных фильтров из H и подавить ложные объекты из S , может быть синтезирован следующим образом [28]:

$$p(x) = \sum_{i=1}^N \alpha_i h_i(x) + \sum_{i=N+1}^{N+M} \alpha_i s_{i-N}(x), \quad (4)$$

где коэффициенты $\{\alpha_i; i = 1, \dots, N + M\}$ выбираются таким образом, чтобы удовлетворить условиям распознавания для объекта из $U = T \cup S$. Здесь мы используем обозначения в векторно-матричной виде: \mathbf{R} — матрица с $N + M$ столбцами и d строками, где каждый столбец является вектором элементом U . Пусть $\mathbf{a} = [\alpha_i; i = 1, \dots, N + M]^T$ — вектор коэффициентов линейной комбинации, тогда можно записать

$$\mathbf{p} = \mathbf{R}\mathbf{a}, \quad (5)$$

где $\mathbf{p} = [p(x); x = 1, \dots, d]$ — дискретная импульсная характеристика композитного фильтра.

Обозначим через

$$\mathbf{u} = \left[\underbrace{1, \dots, 1}_{N \text{ ones}}, \underbrace{0, \dots, 0}_{M \text{ zeros}} \right]^T,$$

отклики системы на объекты обучения, и обозначим через \mathbf{Q} матрицу, столбцы которой являются элементами U . Условия распознавания для каждой версии объекта можно записать как

$$\mathbf{u} = \mathbf{Q}^+ \mathbf{p}, \quad (6)$$

где $^+$ обозначает комплексно-сопряженное транспонирование. Подставив (5) в (6), получим

$$\mathbf{u} = \mathbf{Q}^+ \mathbf{R}\mathbf{a}.$$

Таким образом, решение для \mathbf{a} имеет вид

$$\mathbf{a} = [\mathbf{Q}^+ \mathbf{R}]^{-1} \mathbf{u}. \quad (7)$$

Наконец, подставляя (7) в (5), получаем импульсную характеристику композитного фильтра

$$\mathbf{p} = \mathbf{R}[\mathbf{Q}^+ \mathbf{R}]^{-1} \mathbf{u}. \quad (8)$$

Заметим, что значение корреляционного пика при использовании фильтра, приведённого в выражении (8), ожидается близким к единице для объектов распознавания и близким к нулю для объектов ложного класса.

Коэффициент дискриминации (DC) определяется, как способность фильтра отличать объект от ложных объектов, и задаётся как [28]

$$DC = 1 - \frac{|c^b|^2}{|c^t|^2}, \quad (9)$$

где c^b — это ложный максимум в области фона, а c^t — максимальное значение корреляционного пика, соответствующего объекту. Значение DC близкое к единице указывает на то, что фильтр хорошо отличает объект от любого ложного объекта. Негативные значения DC указывают на то, что фильтр не способен распознать объект. Можно считать, что если значение DC больше заранее установленного порогового значения ($DC > DC_{th}$), то объект считается распознанным с заданным порогом, а иначе объект не распознан.

3. АЛГОРИТМ СЛЕЖЕНИЯ ЗА ОБЪЕКТАМИ НА ОСНОВЕ ЛОКАЛЬНО-АДАПТИВНОЙ КОРРЕЛЯЦИОННОЙ ФИЛЬТРАЦИИ

В этом разделе мы опишем предложенный алгоритм слежения за объектами на основе корреляционной фильтрации. Алгоритм является устойчивым к геометрическим искажениям объектов, неравномерному освещению, частичному перекрытию объектов предметами сцены, сложному фону, а также к аддитивному шуму датчиков.

Алгоритм начинается с процедуры инициализации: выбираются объекты слежения, и для каждого объекта выделяется локальный фрагмент, содержащий объект. Размер фрагмента выбирается в три раза больше прямоугольной рамки, ограничивающей объект. Это сделано для того, чтобы учесть возможные геометрические изменения формы объекта на следующем кадре, а также для улучшения статистических оценок параметров синтезируемого фильтра. Блок-диаграмма алгоритма процедуры инициализации и синтеза банка композитных фильтров для одного объекта показана на Рис. 1. Подробные шаги алгоритма слежения за объектом

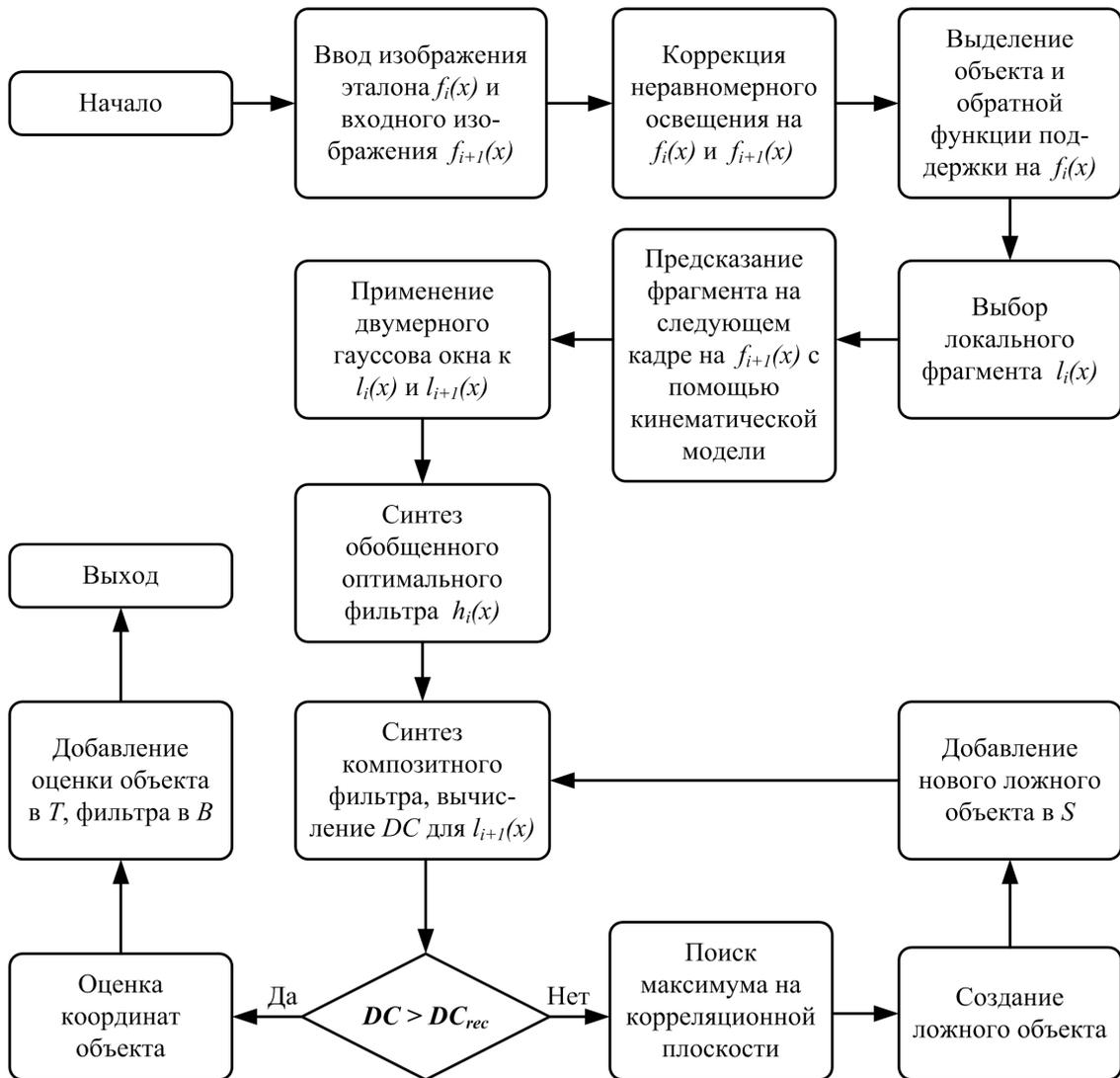


Рис. 1. Блок-диаграмма алгоритма процедуры инициализации и синтеза банка композитных фильтров.

на основе блок-диаграммы описаны ниже.

Шаг 1: Ввод кадров эталона $f_i(x)$ и входной сцены $f_{i+1}(x)$. В самом начале работы объект выбирается в интерактивном режиме с помощью прямоугольной рамки, а на последующих кадрах производится автоматическая оценка функции объекта.

Шаг 2: Коррекция неравномерного освещения изображений $f_i(x)$ и $f_{i+1}(x)$ с помощью гаммы-коррекции [29]:

$$I_o = I^\gamma,$$

где I_o и I — выходное и входное изображения, соответственно. Этот метод корректирует неравномерность освещения и уменьшает локальные эффекты затенения. В областях небольшого размера, таких как фрагмент объекта, освещение становится локально равномерным. Значение параметра γ выбирается равным $1/2$.

Шаг 3: Вычисления оценок объекта $t_i(x)$ и $w_i(x)$ обратной функции поддержки объекта на кадре $f_i(x)$, оценка $t_i(x)$ добавляется в T .

Шаг 4: Выделение локального фрагмента $l_i(x)$, содержащего объект, из $f_i(x)$.

Шаг 5: Предсказание положения следующего фрагмента $l_{i+1}(x)$ на кадре $f_{i+1}(x)$. Координаты фрагмента $l_{i+1}(x)$ предсказываются с помощью модели кинематического движения объекта [30].

Шаг 6: Применение двумерного гауссова окна для уменьшения краевых эффектов и влияния фона на $l_i(x)$ и $l_{i+1}(x)$.

Шаг 7: Синтез обобщенных оптимальных фильтров (3) для обнаружения и оценки положения геометрически измененных версий объекта в локальном фрагменте $l_{i+1}(x)$.

Шаг 8: Синтез композитного линейного фильтра $p_i(x)$ с использованием обобщенных оптимальных фильтров. Обнаружение объекта с помощью композитного фильтра на локальном фрагменте $l_{i+1}(x)$ и вычисление DC . Если DC превышает заданное пороговое значение ($DC > DC_{rec}$), то объект считается успешно распознанным, оцениваются его координаты, оценка $t_{i+1}(x)$ добавляется в T , композитный фильтр $p_i(x)$ добавляется в банк фильтров B , а итеративный процесс синтеза останавливается. В противном случае, вычисляется оценка ложного объекта в точке ложного пика и добавляется в набор S . Производится синтез композитного линейного фильтра $p_i(x)$ с использованием обобщенных оптимальных фильтров и ложных объектов из S , поиск объекта с помощью фильтра на локальном фрагменте $l_{i+1}(x)$. Процедура повторяется итеративно, пока не будет выполнено условие $DC > DC_{rec}$ или количество итераций превысит заданное число. Для экспериментов было выбрано максимальное количество итераций равное 10. Если условия распознавания объекта выполнены, то последний синтезированный композитный фильтр $p_i(x)$ добавляется в банк фильтров B .

Шаг 9: Если на предыдущем шаге синтез композитного фильтры был неуспешным, то рассматриваются следующие варианты: 1) объект находится за пределами локального фрагмента $l_{i+1}(x)$; 2) объект временно закрыт другими предметами сцены или находится за пределами кадра $f_{i+1}(x)$; 3) произошла сильная деформация формы объекта, которая не учитывается при синтезе композитного фильтра. Для решения проблемы (1) алгоритм максимально расширяет зону поиска объекта на текущем кадре, то есть $l_{i+1}(x) = f_{i+1}(x)$, и процедура Шага 8 повторяется. Для решения проблемы (2) алгоритм считывает следующий кадр $f_{i+2}(x)$ и процедура Шагов 2-8 повторяется. Количество считанных кадров зависит от конкретной задачи и является параметром алгоритма. Для экспериментов было выбрано максимальное количество считываний, равное 5. Для решения проблемы (3) алгоритм расширяет спектр возможных версий объекта путем использования предыдущих оценок объекта из T и генерирования дополнительных версий объекта с помощью аффинных преобразований для последующего синтеза композитного фильтра. Одновременно применяется банк успешных фильтров - композитные фильтры из B для распознавания объекта.

4. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ

В этом разделе представлены результаты тестирования предложенного алгоритма слежения за объектами и сравнение его работы с работой известных алгоритмов с точки зрения надежности обнаружения, точности слежения и скорости обработки.

Чтобы оценить качество работы предложенного алгоритма, были проведены эксперименты на видео последовательностях из Object Tracking Benchmark (база данных TB-100 и база данных TB-50) [31]. Эти базы данных содержат последовательности со сложными ситуациями для задачи слежения: изменение освещения сцены (IV), изменение масштаба (SV), частичное перекрытие объектов (OCC), деформация формы (DEF), размытие вследствие движения (MB), быстрое движение (FM), повороты в плоскости (IPR), повороты вне плоскости (OPR), объекты вне поля зрения (OV), сложные фоновые помехи (BC), низкое разрешение (LR). В работе [31] приведены результаты тестирования 33 алгоритмов слеживания за объектами. Наилучшими алгоритмами являются STRUCK и SCM.

Для сравнения с предложенным алгоритмом на этих же базах данных было протестировано четыре известных алгоритма с одинаковым начальным положением объекта: SURF [32], STRUCK [33], SCM [34] и DAC[35]. SURF основан на сопоставлении локальных функций и дескрипторов для обнаружения и распознавания объектов. STRUCK использует машину опорных векторов, чтобы разделить пространство признаков сцены и обнаружить положение объекта с помощью обучающей выборки. SCM объединяет дискриминационный классификатор на основе разреженности и генеративную модель, построенную с помощью гистограмм и учитывающей пространственную информацию для фрагментов каждого кадра. DAC использует обнаружение объектов в фрагментах кадра с перенастраиваемым банком корреляционных фильтров и предсказанием Калмана. Банк композитных фильтров обновляется от кадра к кадру с адаптацией к поворотам и изменениям масштаба объектов.

Были выбраны следующие критерии оценки качества слежения алгоритмов: коэффициент успешного обнаружения (success rate) и точность определения положения объекта [5, 23]. Точность определения положения объекта оценивается ошибкой локализации (LE):

$$LE = \sqrt{(x - \hat{x})^2 + (y - \hat{y})^2},$$

где (x, y) и (\hat{x}, \hat{y}) обозначают точно известные и полученные с помощью алгоритма координаты объекта, соответственно.

Для вычисления коэффициента успешного обнаружения объекта измеряется степень перекрытия между прямоугольной областью нахождения объекта, полученной с помощью алгоритма, и заранее точно известной прямоугольной областью нахождения объекта [5]. Степень перекрытия определяется как

$$S = \frac{\|r_t \cap r_0\|}{\|r_t \cup r_0\|},$$

где r_t — область объекта, найденная с помощью алгоритма, r_0 — точно известная область объекта, \cap и \cup — операции пересечения и объединения, соответственно, $\|\cdot\|$ — количество пикселей в области.

Среднее значение перекрытия (AOS) может быть использовано для оценки коэффициента успешного обнаружения. Алгоритмы слежения также сравниваются по средней ошибке локализации объектов (ALE) на всех видео последовательностях из базы данных. Распространенным способом оценки алгоритма является его инициализация точно известным объектом в первом кадре и последующем вычислении коэффициента успешного обнаружения для всех видео последовательностей базы данных. Такая однопроходная оценка (OPE) также используется в экспериментах.

Протестируем алгоритмы на 50 тестовых видео последовательностях из базы данных ТВ-50. Для каждого алгоритма используются параметры по умолчанию из исходного кода. Результаты тестирования представлены на Рис. 2. Графики показывают коэффициенты успешного

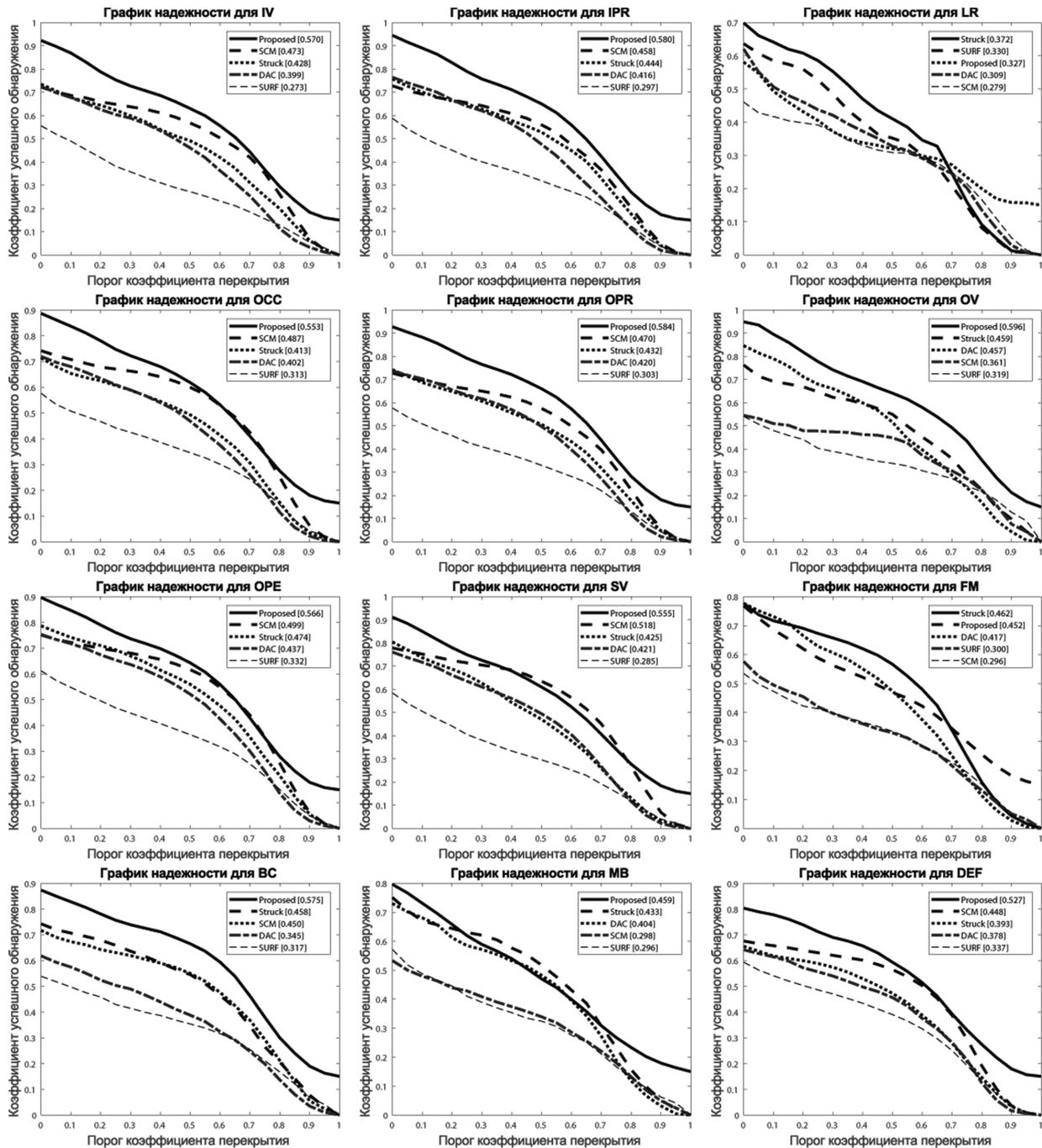


Рис. 2. Коэффициент успешного обнаружения в зависимости от пороговых значений степени перекрытия и площадь области под кривой для алгоритмов STRUCK, SCM, SURF, DAC и предложенного алгоритма (Proposed) для случаев: изменение освещения сцены (IV), повороты в плоскости (IPR), низкое разрешение (LR), частичное перекрытие объектов (OCC), повороты вне плоскости (OPR), объекты вне поля зрения (OV), оценка по всей базе данных (OPE), изменение масштаба (SV), быстрое движение (FM), сложные фоновые помехи (BC), размытие вследствие движения (MB), деформация формы (DEF).

обнаружения в зависимости от пороговых значений степени перекрытия, меняющихся от 0 до 1. Использование одного значения коэффициенты успешного обнаружения при конкретном пороговом значении, например 0.5, для оценки качества алгоритма слежения является недостаточным. Предлагается оценивать качество работы алгоритмов по параметру площади области под кривой (AUC) для каждого графика. Отметим, что предложенный алгоритм (proposed) работает почти для всех тестируемых ситуаций лучше остальных алгоритмов. Наихудший результат из тестируемых алгоритмов показывает SURF из-за плохой динамической модели.

Таблица 1 показывает работу тестируемых алгоритмов с точки зрения надежности обнаружения (коэффициент успешного обнаружения) и точности определения положения объекта (LE) на одной видео последовательности “Баскетбол” из Object Tracking Benchmark (база данных TB-100) [31], искаженной неравномерным освещением и аддитивным шумом с разными значениями отношения энергии сигнала к шуму (SNR).

Таблица 1. Сравнение работы STRUCK, SCM, SURF, DAC и предложенного алгоритма (Proposed) по надежности обнаружения и точности определения положения объекта на видео последовательностях, искаженных неравномерным освещением и аддитивным шумом с различными значениями SNR.

Алгоритмы	Proposed	STRUCK	SCM	SURF	DAC
SNR 50 дБ	100%,3.62	100%,7.34	100%,4.23	71%,23.56	100%,5.54
SNR 25 дБ	100%,3.95	99%,7.57	100%,4.89	70%,24.68	100%,5.83
SNR 15 дБ	100%,4.71	97%,8.18	93%,10.57	64%,26.73	100%,5.94
SNR 10 дБ	100%,4.97	94%,9.92	89%,13.71	60%,30.37	100%,6.47
SNR 5 дБ	100%,5.23	92%,10.30	86%,14.19	54%,33.98	99%,6.93

Из Таблицы 1 видно, что все тестируемые алгоритмы способны обнаруживать объект во всех кадрах сцены при всех уровнях шума при неравномерном освещении сцен. Однако предложенный алгоритм является наиболее робастным, так как его надежность обнаружения и точность вычисления координат объекта являются наилучшими среди тестируемых алгоритмов. Отметим, что алгоритм STRUCK демонстрирует лучшую надежность обнаружения и точность вычисления координат объекта, чем алгоритмы SCM и SURF.

В Таблице 2 приведены средние значение перекрытия (AOS), средняя ошибка вычисления положения центра объекта (ACLE) и среднее время обработки кадра (APT) для всех тестируемых алгоритмов слежения с порогом перекрытия 0.5 на компьютере с процессором Intel Core i5-3340 (3,3 ГГц). Результаты показывают, что предлагаемый алгоритм является самым быстрым и точным с точки зрения средней ошибки определения положения центра объекта.

Также отметим, что при быстром перемещении объекта (FM), деформации формы объекта (DEF), изменении масштаба (SV) предложенный алгоритм работает лучше тестируемых алгоритмов. Однако предложенный алгоритм при выбранных параметрах неэффективен в случае исчезновении объекта на некоторое время из поля зрения (OV). Кроме того, алгоритм можно улучшить для сцен с изменяющимся освещением и частичным перекрытием объектов предметами сцены.

5. ЗАКЛЮЧЕНИЕ

Предложен алгоритм слежения за несколькими объектами, использующий локально-адаптивную обобщенную корреляционную фильтрацию. Для обеспечения устойчивости к геометрическим преобразованиям объектов, частичному перекрытию объектов предметами сцены, неравномерному освещению сцены и наличию аддитивного шума на изображениях сцены алгоритм в каждый момент времени строит композитные фильтры, которые адаптируются к

Таблица 2. Сравнение работы STRUCK, SCM, SURF, DAC и предложенного алгоритма (Proposed) с точки зрения среднего значения перекрытия (AOS), средней ошибки определения положения (ALE) и среднего времени обработки кадра (APT).

Алгоритм	Proposed	STRUCK	SCM	SURF	DAC
BC	49.1	59.3	61.3	37.4	41.3
DEF	54.8	52.4	51.5	25.8	50.4
FM	61.2	55.6	42.8	41.6	43.7
IPR	59.2	57.0	51.8	39.7	42.6
IV	53.7	59.0	61.1	37.3	44.7
LR	48.3	59.1	61.7	23.0	39.7
MB	58.1	59.9	45.2	45.4	48.3
OCC	50.6	55.9	56.8	36.0	46.2
OPR	55.3	57.3	57.0	34.8	50.1
OV	39.4	58.9	56.4	46.7	39.8
SV	57.9	57.8	55.8	33.0	52.1
All	59.2	57.5	54.4	35.2	43.1
APT	0.1749	0.2894	0.3122	0.1668	0.3255
ALE	49.2	61.5	64.8	276.6	69.4

текущим изменениям сцены с помощью итеративного обучения. Предложенная система использует схему предсказания, построенную на основе кинематической модели движения объектов. Работа предложенного алгоритма сравнивается с работой известных алгоритмов видео слежения на известных базах данных с точки зрения надежности обнаружения, точности вычисления траектории объектов и скорости обработки. Результаты тестирования показывают, что предложенный алгоритм является быстрым, более точным и надежным на большинстве видео последовательностей из базы данных, чем известные алгоритмы слежения.

СПИСОК ЛИТЕРАТУРЫ

1. Karasulu B., Korukoglu S. Performance Evaluation Software: Moving Object Detection and Tracking in Videos. New York : Springer, 2013.
2. Talmale S., Janwe N. Object tracking in images and videos // International Journal Of Engineering And Computer Science. 2016. Vol. 5(1). P. 15482—15486.
3. Vijaya Kumar B. V. K., Fernandez J. A., Rodriguez A., Boddeti V. N. Recent advances in correlation filter theory and application // Proc. SPIE. Vol. 9094. 2014. P. 909404—13.
4. Kumar B. V. K. V., Mahalanobis A., Juday R. D. Correlation Pattern Recognition. New York, NY, USA : Cambridge University Press, 2005.
5. Smeulders A. W. M., Chu D. M., Cucchiara R. et al. Visual tracking: An experimental survey // IEEE Trans. Pattern Anal. Mach. Intell. 2014. Vol. 36, no. 7. P. 1442—1468.
6. Akin O., Erdem E., Erdem A., Mikolajczyk K. Deformable part-based tracking by coupled global and local correlation filters // Journal of Visual Communication and Image Representation. 2016. Vol. 38, no. Supplement C. P. 763—774.
7. Ali A., Jalil A., Ahmed J. et al. Correlation, kalman filter and adaptive fast mean shift based heuristic approach for robust visual tracking // Signal, Image and Video Processing. 2015. Vol. 9, no. 7. P. 1567—1585.
8. Henriques J. F., Caseiro R., Martins P., Batista J. High-speed tracking with kernelized correlation filters // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2015. Vol. 37, no. 3. P. 583—596.

9. Li X., Liu Q., He Z. et al. A multi-view model for visual tracking via correlation filters // Knowledge-Based Systems. 2016. Vol. 113, no. Supplement C. P. 88—99.
10. Li Z., Yang J., Zha J. et al. Online visual tracking via correlation filter with convolutional networks // 2016 Visual Communications and Image Processing (VCIP). 2016. P. 1—4.
11. Liu F., Zhou T., Fu K., Yang J. Robust visual tracking via constrained correlation filter coding // Pattern Recognition Letters. 2016. Vol. 84, no. Supplement C. P. 163—169.
12. Liu F., Zhou T., Yang J. Geometric affine transformation estimation via correlation filter for visual tracking // Neurocomputing. 2016. Vol. 214, no. Supplement C. P. 109—120.
13. Jeong S., Kim G., Lee S. Effective visual tracking using multi-block and scale space based on kernelized correlation filters // Sensors. 2017. Vol. 17, no. 3. P. 433.
14. Bibi A., Mueller M., Ghanem B. Target response adaptation for correlation filter tracking // Computer Vision – ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part VI. Cham : Springer International Publishing, 2016. P. 419—433.
15. Gundogdu E., Ozkan H., Alatan A. A. Extending correlation filter-based visual tracking by tree-structured ensemble and spatial windowing // IEEE Transactions on Image Processing. 2017. Vol. 26, no. 11. P. 5270—5283.
16. Ontiveros-Gallardo S. E., Kober V. Objects tracking with adaptive correlation filters and kalman filtering // Proc. SPIE. Vol. 9598. 2015. P. 95980X—8.
17. Díaz-Ramírez V. H., Picos K., Kober V. Object tracking in nonuniform illumination using space-variant correlation filters // Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications: 18th Iberoamerican Congress, CIARP 2013, Havana, Cuba, November 20-23, 2013, Proceedings, Part II. Berlin, Heidelberg : Springer Berlin Heidelberg, 2013. P. 455—462.
18. Diaz-Ramirez V. H., Picos K., Kober V. Target tracking in nonuniform illumination conditions using locally adaptive correlation filters // Optics Communications. 2014. Vol. 323. P. 32—43.
19. Diaz-Ramirez V. H., Contreras V., Kober V., Picos K. Real-time tracking of multiple objects using adaptive correlation filters with complex constraints // Optics Communications. 2013. Vol. 309. P. 265—278.
20. Leopoldo N. Gaxiola Victor H. Diaz-Ramirez J. J. T. Target tracking using interest point detection and correlation filtering // Proc. SPIE. Vol. 9217. 2014. P. 9217—8.
21. Gaxiola L. N., Díaz-Ramírez V. H., Tapia J. J. et al. Robust face tracking with locally-adaptive correlation filtering // Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications: 19th Iberoamerican Congress, CIARP 2014, Puerto Vallarta, Mexico, November 2-5, 2014. Proceedings. Cham : Springer International Publishing, 2014. P. 925—932.
22. Cuevas A., Diaz-Ramirez V. H., Kober V., Trujillo L. Facial recognition using composite correlation filters designed with multiobjective combinatorial optimization // Proc. SPIE. Vol. 9217. 2014. P. 921710—8.
23. Gaxiola L. N., Diaz-Ramirez V. H., Tapia J. J., Garcia-Martinez P. Target tracking with dynamically adaptive correlation // Optics Communications. 2016. Vol. 365. P. 140—149.
24. Ruchay A., Kober V. A correlation-based algorithm for recognition and tracking of partially occluded objects // Proc. SPIE. Vol. 9971. 2016. P. 99712R—9.
25. Ruchay A., Kober V., Gonzalez-Fraga J. A. Reliable recognition of partially occluded objects with correlation filters // Mathematical Problems in Engineering. 2018. Vol. 2018. P. 8284123.
26. Aguilar-Gonzalez P. M., Kober V. Design of correlation filters for pattern recognition with disjoint reference image // Optical Engineering. 2011. Vol. 50, no. 11. P. 117201.
27. Aguilar-Gonzalez P. M., Kober V. Design of correlation filters for pattern recognition using a noisy reference // Optics Communications. 2012. Vol. 285, no. 5. P. 574—583.

28. Ramos-Michel E. M., Kober V. Adaptive composite filters for pattern recognition in linearly degraded and noisy scenes // *Optical Engineering*. 2008. Vol. 47, no. 4. P. 047204–7.
29. Lee P.-M., Chen H.-Y. Adjustable gamma correction circuit for tft lcd // *Proceedings - IEEE International Symposium on Circuits and Systems*. 2005. 01. P. 780—783.
30. Gupta P. *Comprehensive Business Statistics*. Laxmi Publications, 2005.
31. Wu Y., Lim J., Yang M. H. Object tracking benchmark // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2015. Vol. 37, no. 9. P. 1834—1848.
32. Al-asadi T., Obaid A. Object detection and recognition by using enhanced speeded up robust feature // *International Journal of Computer Science and Network Security*. 2016. Vol. 16(4). P. 66—71.
33. Torr P. H. S., Hare S., Saffari A. Struck: Structured output tracking with kernels // *2011 IEEE International Conference on Computer Vision (ICCV 2011)*. 2011. Vol. 00. P. 263—270.
34. Zhong W. Robust object tracking via sparsity-based collaborative model // *Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. CVPR '12. Washington, DC, USA : IEEE Computer Society, 2012. P. 1838—1845.
35. Ontiveros-Gallardo S. E., Kober V. Correlation-based tracking using tunable training and kalman prediction // *Proc. SPIE*. Vol. 9971. 2016. P. 997129–9.

Tracking of multiple objects with locally adaptive generalized optimum correlation filters

Kober V.I., Ruchay A.N., Karnaukhov V.N.

An algorithm for tracking multiple objects using locally-adaptive generalized correlation filtering is proposed. The algorithm is invariant to geometric transformations of objects, partial occlusion of objects, nonuniform illumination, and additive noise. The proposed system utilizes generalized optimal correlation filters and a prediction scheme based on the kinematic model of objects motion. With the help of iterative training the system can be adapted to current scene changes. The performance of the proposed algorithm is compared with that of the state-of-the-art visual tracking algorithms on known databases in terms of quality metrics and processing time.

KEYWORDS: tracking; adaptive filter; correlation filter.