Оглавление

[Вступление 1](#_Toc389192356)

[1. Аналитический раздел 3](#_Toc389192357)

[1.1. Обзор подходов к отслеживанию объектов в видеопотоке 3](#_Toc389192358)

[1.2. Основные принципы вероятностного отслеживания 8](#_Toc389192359)

[1.3. Фильтр частиц 11](#_Toc389192360)

[Выводы 12](#_Toc389192361)

[2. Конструкторский раздел 13](#_Toc389192362)

[2.1. Алгоритм воспроизведения условной плотности 13](#_Toc389192363)

[2.2. Выбор способа представления объекта 14](#_Toc389192364)

[2.3. Модель динамики объекта 15](#_Toc389192365)

[2.4. Описание внешних признаков объекта 15](#_Toc389192366)

[2.5. Процесс измерения состояния объекта 16](#_Toc389192367)

[2.6. Алгоритм отслеживания 17](#_Toc389192368)

[2.7. Адаптация стохастической компоненты модели динамики 18](#_Toc389192369)

[Выводы 20](#_Toc389192370)

[3. Технологический раздел 21](#_Toc389192371)

[4. Исследовательский раздел 22](#_Toc389192372)

[4.1. Формирование набора данных 22](#_Toc389192373)

[4.2. Метрики, используемые для оценки работы метода 22](#_Toc389192374)

[4.3. Результаты экспериментов 23](#_Toc389192375)

[4.4. Исследование времени обнаружения объекта с момента инициализации 26](#_Toc389192376)

[Заключение 28](#_Toc389192377)

[Список литературы 29](#_Toc389192378)

[Приложение А 30](#_Toc389192379)

[Классификация подходов к визуальному отслеживанию объектов 30](#_Toc389192380)

# Вступление

Визуальное отслеживание заключается в последовательном определении местоположения целевого объекта на каждом кадре видеопотока. Эта задача находит практическое применение во многих областях, так или иначе связанных с обработкой видеозаписей для получения из них некоторой информации. Примерами систем, в которых применяются технологии отслеживания, являются системы человеко-машинного взаимодействия, видеонаблюдения, анализа спортивных матчей, появившиеся в недавнее время системы «умного дома», системы дополненной реальности и многие-многие другие.

Несмотря на значительное количество научных работ и исследований, проводимых в данной области, точное и устойчивое к ошибкам отслеживание объектов на видео остается сложной проблемой. Небольшие размеры объектов, меняющийся фон, резкие ускорения и смены траектории движения объектов, частичные или полные перекрытия, изменение масштаба отслеживаемой цели при ее приближении или удалении от камеры являются основными трудностями на пути к построению точной траектории перемещения объекта на кадрах видеозаписи.

Целью магистерской работы является разработка метода отслеживания объектов в видеопотоке, позволяющего справляться с указанными проблемами. Для достижения поставленной цели необходимо в первую очередь проанализировать существующие методы отслеживания, определить их сильные и слабые стороны и выбрать из них тот, который будет положен в основу разрабатываемого метода. Также необходимо сформулировать математическое описание предлагаемого метода и разработать на его основе алгоритм отслеживания. Для проверки работоспособности предложенного алгоритма требуется разработать реализующее его программное обеспечение и сформировать наборы данных для проведения исследования. По результатам проведенного исследования необходимо сделать вывод о работоспособности метода и накладываемых на него ограничениях.

Работа построена следующим образом. В аналитическом разделе приводится обзор существующих методов отслеживания с указанием их достоинств и недостатков, принимается решение о выборе метода, на основе которого будет вестись дальнейшая разработка. Здесь же приводится математическое описание выбранного метода. В конструкторском разделе описывается разрабатываемый алгоритм отслеживания. Технологический раздел посвящен описанию разработки программного обеспечения, реализующего данный алгоритм, выбору языка и среды программирования, используемых библиотек. В исследовательском разделе описываются условия, при которых проверяется разработанный метод, приводятся метрики оценки его качества. По результатам исследований делается вывод о применимости разработанного метода в тех или иных задачах отслеживания, его достоинствах и накладываемых на него ограничениях, предлагаются направления дальнейшего исследования.

# Аналитический раздел

В данном разделе приводится классификация существующих методов отслеживания, описываются их достоинства и недостатки, обосновывается выбор метода для дальнейшей разработки и дается его формальное описание.

## Обзор подходов к отслеживанию объектов в видеопотоке

Целью визуального отслеживания является построение траектории движения объекта, то есть последовательности его положений на каждом кадре видеопотока. В качестве дополнительной информации алгоритм отслеживания (трекер, англ. tracker) может предоставлять информацию о занимаемой объектом области на кадре. Задачи распознавания области объекта на кадре и сопоставления данных областей на соседних кадрах могут решаться как раздельно, так и совместно. В связи с этим можно выделить два основных подхода к отслеживанию объектов.

В первом подходе ключевую роль играют алгоритмы распознавания . С их помощью определяются новые объекты, появляющиеся в сцене, и инициализируются трекеры. В этом случае процесс отслеживания сводится к задаче сопоставления объектов, распознанных на текущем кадре, с объектами, выявленными на предыдущем кадре. Такой подход реализует отслеживание на основе распознавания.

В качестве примера алгоритма, реализующего данный подход, можно привести алгоритм AdaBoost, описанный в работе [5]. Для представления объекта авторы работы используют признаки Хаара (англ. Haar-like features), гистограммы направленных градиентов (англ. Histogram of Oriented Gradients, HOG) и локальные бинарные шаблоны (англ. Local Binary Patterns, LBP). Задача соотнесения объектов на соседних кадрах решается с помощью процесса сопоставления с эталоном (англ. template matching) и бинарного классификатора.

Использование алгоритмов распознавания в качестве основы для процесса отслеживания позволяет повысить качество получаемых результатов за счет предоставления трекеру достаточно точного описания занимаемой объектом области. С другой стороны, для успешного определения области объекта на кадре необходимо, чтобы объект имел отличительные внешние характеристики (форма, цвет, текстура и т. п.), а для соотнесения выявленных областей требуется, чтобы объект не претерпевал значительных изменений на соседних кадрах. Таким образом, алгоритмы, реализующие данный подход, неэффективны в тех случаях, когда нельзя заранее предсказать визуальные характеристики цели, например, при быстрых перемещениях объекта, когда его внешний вид также меняется достаточно быстро. Данный класс алгоритмов не подходит также для отслеживания целей небольших размеров и целей, не имеющих достаточных отличительных признаков.

Второй поход к визуальному отслеживанию основан на совмещении процессов выделения областей объекта на кадре и их соотнесения между соседними кадрами. Основная идея заключается в итеративном обновлении положения объекта и занимаемой им области на каждом кадре, то есть область объекта вычисляется на основе результатов отслеживания, полученных на предыдущих кадрах. Данный подход реализует распознавание на основе отслеживания. Его отличительной особенностью является описание множества возможных состояний объекта с помощью некоторой функции, анализ которой позволяет получить информацию о его текущем положении.

Для дальнейшего изучения был выбран второй подход, так как он позволяет учитывать динамику объекта за счет использования всей информации о положении и области объекта, полученной с начала отслеживания, а не только с предыдущего кадра. Также данный подход является менее затратным с вычислительной точки зрения, поскольку в большинстве своем алгоритмы, реализующие распознавание на основе отслеживания, используют более слабую модель представления объекта, позволяя избавиться от затрат на распознавание объекта и снимая требование наличия у объекта отличительных внешних характеристик.

Алгоритмы, реализующие распознавание на основе отслеживания, делятся на две группы: детерминированные и стохастические (вероятностные) алгоритмы. Наиболее известными представителями класса детерминированных алгоритмов отслеживания являются сдвиг среднего (англ. Mean Shift) и непрерывно адаптивный сдвиг среднего (англ. Continuously Adaptive Mean-shift, CAM-shift). Их основная идея заключается в задании для каждого кадра весовой функции, такой, чтобы искомое положение объекта соотносилось с ее минимум или максимумом, и сведении задачи отслеживания к оптимизации данной функции. Алгоритмы сдвига среднего основаны на следующем принципе: на каждом кадре ищется область фиксированной формы с переменным размером, цветовое содержание которой наиболее близко к эталонной цветовой модели. Поиск начинается с положения, определенного на предыдущем кадре, и продолжается до тех пор, пока не будет найден минимум функции, задающей расстояние до эталонной цветовой гистограммы. Данный метод показывает достаточно высокие результаты, кроме тех случаев, когда объект находится на фоне, близком к нему по цвету, а также если он на некоторое время полностью перекрыт другими элементами сцены. Общий недостаток детерминированных подходов заключается в низкой способности к восстановлению после потери объекта, поскольку поиск нового положения всегда начинается с последней точки, которая может быть определена некорректно.

Стохастические (вероятностные) алгоритмы отслеживания основаны на теории байесовского оценивания. В них задача отслеживания сводится к итеративной оценке функции плотности распределения вероятности вектора состояния объекта, производимой на основе измерения его значений на последовательных кадрах. Под измерением будем понимать процесс получения оценки возможного состояния объекта на основе информации, извлекаемой из текущего кадра. Результат процесса измерения является наблюдением текущего состояния объекта. Вектор состояния объекта может описывать его положение, размер, цвет, скорость, ускорение и другие характеристики. Вероятностные методы отслеживания позволяют учитывать погрешности, получаемые в ходе измерения значения вектора состояния, а также шумовые воздействия, которым подвержен данный процесс. Помимо этого учитываются стохастические ошибки модели динамики объекта, позволяющие моделировать движение с резкими изменениями направления и скорости. Однако для получения фактической оценки состояния объекта из получаемых наблюдений необходимо избавиться от данных погрешностей. Для этого часто применяют процесс фильтрации.

Широко известным представителем данного класса методов отслеживания является фильтр Кальмана (англ. Kalman Filter). Он основан на предположении, что функция плотности распределения вероятности вектора состояний является Гауссовой, что позволяет получить аналитическое решение ее оценки. К сожалению, данная гипотеза верна далеко не во многих случаях. Для решения этой проблемы были разработаны модификации данного метода: расширенный фильтр Кальмана (англ., Extended Kalman Filter, EKF), применяющий процесс линеаризации, и ансцентный фильтр Кальмана[[1]](#footnote-1) (англ. Unscented Kalman Filter, UKF), использующий детерминированные выборки. Они относятся к категории параметрических методов, работающих с нелинейными функциями плотности распределения [9]. Однако в большинстве своем эти решения показывают удовлетворительные результаты только при задании строгих ограничений на область применения алгоритма. Обычно они требуют наличия большого количества параметров, настройка которых поможет адаптировать алгоритм для решения конкретной задачи. Несмотря на возможность обработки функций плотности распределения вероятности, не имеющих аналитической оценки, параметрические методы вероятностного отслеживания всегда пытаются свести их к Гауссовому распределению. Но если реальная плотность распределения вероятности состояния объекта имеет, например, бимодальный или сильно ассиметричный вид, она не может быть точно описана нормальным распределением. Для таких случаев больше подходят непараметрические методы отслеживания, позволяющие провести аппроксимацию искомой плотности распределения [1].

Если пространство состояний объекта может быть разбито на ячеек, для аппроксимации функции плотности распределения вероятности вектора состояния объекта может быть использован сеточный метод (англ. approximate grid-based method), суть которого заключается в вычислении весового коэффициента для центра каждой ячейки. При этом сетка должна быть достаточно плотной для получения хорошего приближения к непрерывному пространству состояний. Поэтому при увеличении размерности пространства состояний резко возрастает и вычислительная стоимость сеточной аппроксимации. Также данный метод требует заранее определенную сетку, поэтому пространство состояний не может иметь неравномерное разбиение для получения более высокого разрешения в областях с высокой плотностью распределения вероятности, если только эти области не известны заранее [1].

К классу непараметрических методов отслеживания относятся также методы Монте-Карло, позволяющие работать с несколькими гипотезами одновременно, что обуславливает их естественную способность адаптироваться к изменениям, происходящим с отслеживаемым объектом. Среди данной группы методов можно выделить фильтр частиц (англ. particle filter), аппроксимирующий функцию плотности распределения вероятности объекта набором взвешенных частиц.

Иллюстрация изложенной классификации методов визуального отслеживания представлена в приложении А. Проведенный анализ показал, что распознавание на основе отслеживания позволяет получить более точную оценку динамики объекта, поскольку методы данного класса используют всю историю отслеживания для вычисления следующего состояния объекта, а не только данные с предыдущего шага. Вероятностные методы отслеживания, в свою очередь, повышают точность этой оценки за счет принятия во внимание случайных погрешностей модели динамики объекта и шумовых воздействий на процесс измерения его состояния. Поскольку в большинстве реальных задач отслеживания плотность распределения вектора состояния объекта не имеет аналитического представления, необходимо применять методы, способные работать с распределениями, которые не могут быть сведены к Гауссовому без значительной потери точности. Такие методы используют различные способы аппроксимации искомой функции плотности распределения. В данной работе предпочтение отдано фильтру частиц, поскольку он имеет ряд значительных преимуществ. Во-первых, он не использует никаких предположений о характере функции плотности распределения вероятности состояния объекта. Во-вторых, он позволяет анализировать только часть кадра, в отличие от сеточной аппроксимации, при которой на ячейки разбивается вся область полученного с камеры изображения, что позволяет снизить накладные расходы на процесс измерения состояния объекта. В-третьих, увеличение размерности пространства состояний, то есть добавление новых характеристик в вектор состояния объекта, не влечет за собой резкого повышения вычислительной стоимости. Еще одним его достоинством по сравнению с сеточной аппроксимацией является способность повышать концентрацию частиц в областях с высокой плотностью распределения вероятности без использования каких-либо априорных предположений о ее характере. В заключение следует отметить, что фильтр частиц достаточно прост с точки зрения реализации, что позволяет снизить затраты на разработку программного обеспечения.

Среди недостатков фильтра частиц следует выделить необходимость его настройки под конкретную задачу отслеживания. С одной стороны, это повышает гибкость и универсальность метода. Но при этом задача выбора используемых распределений вероятности и их параметров является далеко не тривиальной и требует в некоторых случаях проведения отдельного исследования. Низкая вычислительная стоимость фильтра частиц обусловлена во многом использованием слабого низкоуровневого представления объекта, чаще всего, с помощью цветового распределения, что предполагает наложение ряда ограничений, особенно в тех случаях, когда целевой объект не имеет характерных отличительных внешних признаков.

Несмотря на указанные недостатки, в основу дальнейшей разработки положен аппроксимация функции плотности распределения вероятности состояния объекта с помощью фильтра частиц. Для более глубоко изучения данного метода требуется вначале описать общий принцип вероятностного отслеживания, на котором базируется фильтр частиц.

## Основные принципы вероятностного отслеживания

Как было отмечено ранее, вероятностные методы визуального отслеживания основаны на байесовском подходе, который заключается в задании состояния объекта с помощью функции плотности распределения его вероятности. Данная функция характеризует текущую степень знания о состоянии объекта и строится на основе всей полученной в ходе отслеживания информации.

Задачу отслеживания можно сформулировать с помощью следующего математического представления пространства состояний объекта:

(1.1)

(1.2)

Уравнение  описывает изменение состояния объекта при переходе от кадра к кадру и задает модель динамики объекта. Фактически состояние зависит от предыдущего состояния объекта и случайной ошибки , представляющей собой погрешность, получаемую при обновлении вектора состояния объекта. Поскольку ошибка является случайной величиной с известной статистикой, уравнение  неявно задает функцию плотности распределения вероятности . Уравнение  описывает процесс измерения, результатом которого является некоторая величина , называемая наблюдением. Наблюдение зависит от текущего состояния и случайной ошибки , описывающей шумовое воздействие, оказываемое на процесс измерения состояния объекта. Аналогично ), поскольку является случайной величиной, уравнение  неявно задает функцию правдоподобия .

С точки зрения байесовского подхода проблема отслеживания заключается в рекурсивном вычислении некоторой степени доверия к состоянию объекта на кадре с учетом всех имеющихся к текущему моменту наблюдений. Таким образом, для каждого кадра необходимо вычислить функцию плотности распределения вероятности состояния объекта ), называемую постериорной (англ. posterior). Для ее вычисления необходимы результаты, полученные на предыдущем шаге (плотность распределения вероятности ), а также априорная плотность распределения вероятности смены состояния (англ. state transition prior) и функция правдоподобия (рис.Рис. 1.1).

Вычисление плотности распределения вероятности состояния объекта состоит из двух этапов: прогноза и коррекции (рис. Рис. 1.2). Полагая известным значение искомой функции на предыдущем кадре, с помощью уравнения Чепмена-Колмогорова  можно вычислить априорную функцию плотности распределения вероятности состояния объекта на кадре :

(1.3)

|  |
| --- |
| idef0.jpg |
| Рис. 1.1. Вычисление плотности распределения вероятности состояния объекта на одном кадре |

|  |
| --- |
| idef0_1.jpg |
| Рис. 1.2. Этапы вычисления плотности распределения вероятности состояния объекта на кадре : предсказание и коррекция. |

В этом заключается этап предсказания. На кадре становится известным наблюдение , которое может быть использовано для обновления априорной функции по правилу Байеса :

(1.4)

где — нормирующий множитель, зависящий от функции правдоподобия.

Рекурсивные соотношения  и  формируют базис для нахождения оптимального байесовского решения. Такой подход является концептуальным с той точки зрения, что в общем случае данное решение не может быть получено аналитическим путем. Примером частного случая, когда аналитическое решение поставленной задачи все-таки существует, является упомянутый в предыдущей главе фильтр Кальмана. В более сложных случаях искомое решение можно получить с помощью аппроксимации. Одним из возможных способов для этого является фильтр частиц [1].

## Фильтр частиц

В основе метода лежит аппроксимация функции плотности распределения вероятности состояния объекта набором частиц, поведение которых регулируется с помощью их весов. Фактически, частица является представлением некоторого возможного состояния объекта. Множество   задает приближенную функцию плотности распределения вероятности для состояния при заданном наборе наблюдений .

(1.5)

где – количество частиц, — вес частицы , задаваемый с помощью функции правдоподобия. Вес частицы показывает вероятность, с которой объект примет состояние, описываемое данной частицей. Веса частиц нормализованы: . Тогда искомая постериорная плотность распределения вероятности состояния объекта может быть представлена как (1.6):

(1.6)

где — дельта функция Дирака.

Теперь задача вычисления плотности распределения вероятности состояния объекта на кадре сводится к построению множества на основе полученного на предыдущем шаге множества . Для этого необходимо сформировать новый набор частиц, производя выборку из согласно некоторому «прогнозному» распределению (англ. proposal distribution) , зависящему в общем случае от предыдущего состояния объекта и получаемого на текущем кадре наблюдения, т.е:

(1.7)

Выбор подходящего распределения зависит от особенностей решаемой задачи и способен оказать значительное влияние на получаемые результаты. Наиболее часто встречается случай, когда совпадает с плотностью распределения вероятности смены состояния , но известны и другие варианты, как, например, в работе [2].

После формирования нового набора частиц необходимо пересчитать их весовые коэффициенты согласно выражению (1.8) и нормализовать их.

(1.8)

При новые вычисленные весовые коэффициенты пропорциональны функции правдоподобия .

Таким образом, выражение (1.8) описывает процесс коррекции (1.4) в терминах фильтра частиц.

## Выводы

В данном разделе была проведена классификация существующих подходов к визуальному отслеживанию объектов, проанализированы их достоинства и недостатки, обоснован выбор фильтра частиц в качестве основы для дальнейшей разработки. Кроме этого, было дано аналитическое описание основных принципов вероятностного отслеживания и описана и их интерпретация в терминах фильтра частиц.

# Конструкторский раздел

В данном разделе приводится алгоритмическая основа разрабатываемого метода, описывается способ представления объекта на кадре и процесс измерения его состояния, задается модель динамики объекта. Также описывается предлагаемая модификация стандартного алгоритма.

## Алгоритм воспроизведения условной плотности

Основные принципы работы фильтра частиц можно рассмотреть на примере алгоритма воспроизведения условной плотности (англ. Conditional Density Propagation, Condensation). Он является базовым алгоритмом фильтра частиц, на основе которого разрабатываются различные модификации для его адаптации под конкретные условия отслеживания. Основная идея алгоритма заключается в построении множества частиц на основе множества , полученного на предыдущем шаге.

На каждой итерации алгоритма полагаем множество пустым. Алгоритм включает в себя 3 этапа:

* перевыборка (англ. resampling);
* прогноз;
* коррекция (обновление весов).

В алгоритме воспроизведения условной плотности в качестве прогнозного распределения , согласно которому производится выборка новых частиц (1.7), используется плотность распределения вероятности смены состояния . Это означает, что частица из множества попадает в новый набор с вероятностью, равной ее весу. В результате будет получен набор из частиц, в котором несколько раз может встретиться одна и та же частица из предыдущего набора.

На этапе прогноза каждая частица полученного множества обновляется согласно уравнению динамики :

(2.1)

Этот шаг алгоритма соответствует вычислению априорной функции плотности распределения вероятности.

Выражение (1.8) описывает этап коррекции, который заключается в обновлении весов полученных частиц. При этом, так как в качестве прогнозного распределения используется , веса частиц вычисляются пропорционально выбранной функции правдоподобия. После обновления необходимо провести нормализацию весов, так, чтобы их общая сумма была равна единице (2.2).

(2.2)

В результате каждой итерации будет получено искомое множество частиц . Оценка состояния объекта на текущем шаге может быть вычислена как «средняя» частица полученного множества (2.3):

(2.3)

Так выглядит алгоритм воспроизведения условной плотности в общем виде. Чтобы реализовать его в рамках конкретной задачи, необходимо выбрать способ представления объекта и описания его внешних характеристик, модель его динамики, описать процесс измерения состояния объекта и задать функцию правдоподобия.

## Выбор способа представления объекта

В задачах визуального отслеживания применяются следующие способы представления объекта [10]:

* набор точек (обычно применяется в ситуациях, когда объект занимает небольшую область кадра);
* геометрический примитив, например, прямоугольник или эллипс (чаще всего применяется для простых объектов фиксированной формы, но может использоваться и для объектов с изменяющейся формой);
* силуэт или контур (подходит для объектов, имеющих сложную изменяющуюся форму);
* шарнирная (сочлененная) модель (англ. articulated shape model) (например, представление человека в виде соединенных суставами частей тела);
* скелетная модель (чаще всего используется для распознавания объекта).

Алгоритмы, реализующие фильтры частиц, обычно используют для представления объекта геометрические примитивы [2, 4, 6, 7], иногда наборы точек [3]. В данной работе объект представлен с помощью описывающего прямоугольника, так как такой способ позволяет оценить размер области, занимаемой объектом на кадре, а также удобен для использования с моделью цветового распределения, обычно применяемой в вероятностных методах отслеживания для описания внешних признаков объекта.

Вектор состояния объекта (2.4) включает в себя координаты верхнего левого угла прямоугольника, его ширину и высоту . Эти характеристики задают статическую часть вектора состояния объекта. Вторая компонента вектора представляет его динамическую часть, описывая производные статических элементов из .

(.)

## Модель динамики объекта

Уравнение динамики объекта (1.1) описывает изменение состояния объекта при переходе от кадра к кадру виеодеозаписи и используется для предсказания возможного состояния объекта на следующем кадре. Обычно в качестве уравнения динамики выбирают модель случайных блужданий (англ. Random Walk Model) [6], модель динамики первого порядка [4, 7] или или авторегрессионную модель второго порядка [2, 8].

В данной работе применяется модель динамики первого порядка, позволяющяя учитывать скорость движения объекта при предсказании его следующего состояния. В рамках этой модели уравнение  будет иметь вид :

(2.5)

где — единичная матрица размера , — шаг по времени. Так как скорость измеряется (в пиксель/кадр) между последовательными кадрами, . — аддитивный изотропный Гауссов шум с нулевым математическим ожиданием и дисперсией, задаваемой вектором стандартных отклонений , в котором также выделяются статическая и динамическая компоненты.

(2.6)

* 1. Описание внешних признаков объекта

Вероятностные методы отслеживания, и в частности фильтр частиц, используют для описания внешних признаков объекта модель цветового распределения (англ. Color Distribution Model). Описание объекта с помощью цветового распределения позволяет повысить ошибкоустойчивость получаемых наблюдений для подвижных объектов, учитывая их возможное вращение и частичные перекрытия [7].

Дискретным представлением цветового распределения является гистограмма с карманами. Она может быть построена как в цветовом пространстве RGB, так и в HSV, причем в последнем случае снижается чувствительность к изменениям освещения, поскольку в пространстве HSV значение интенсивности (Value) отделено от цветовых составляющих: оттенка (Hue) и насыщенности (Saturation), и его можно не учитывать при построении гистограммы. Цветовая гистограмма строится по области, ограниченной прямоугольником, представляющим объект.

Цветовое распределение для прямоугольной области с левым верхним углом в точке вычисляется как :

(2.7)

где — общее количество пикселей в области, — дельта-функция Кронекера, — номер кармана гистограммы, — функция, сопоставляющая цвету в точке области номер соответствующего кармана гистограммы.

## Процесс измерения состояния объекта

Под процессом измерения понимается получение информации о текущем состоянии объекта на основе данных, полученных после обработки очередного кадра. Результат процесса измерения будем называть наблюдением.

Предположим, что известна эталонная цветовая гистограмма объекта . Тогда процесс измерения можно определить как сравнение эталонной гистограммы с гистограммами, построенными по областям кадра, задаваемым каждой из частиц. За текущее наблюдение будет принята та область, цветовая гистограмма которой наиболее близка к эталонной.

Различие между гистограммами может быть выражено такими количественными оценками, как корреляция, пересечение, расстояние Пирсона (), расстояние (или коэффициент) Бхаттачария (англ. Bhattacharyya distance) [4]. Наиболее часто применяемой оценкой является расстояние Бхаттачария. Для двух гистограмм и оно определяется следующим образом :

(2.8)

Чем меньше расстояние , тем ближе цветовая гистограмма , заданная частицей , к эталонной гистограмме , то есть тем с большей долей вероятности частица описывает реальное состояние объекта. Соответственно, данной частице надо назначить более высокий вес, чтобы она попала в набор частиц для следующего кадра. Тогда функция правдоподобия , может задавать нормальное распределение величины (2.8) для вычисления веса частицы (2.9):

(2.9)

## Алгоритм отслеживания

С учетом описанных параметров можно конкретизовать алгоритм воспроизведения условной плотности следующим образом. Пусть задан начальный набор частиц . Для всех кадров видеопоследовательности требуется выполнить следующие действия:

1. Инициализировать набор частиц на кадре пустым множеством: .
2. Для каждой частицы ычислить интегральные веса по формуле (2.10).

(2.10)

1. Определить экземпляр выборки . Для этого выбрать случайное число и вычислить . Добавить в . Повторять данную операцию, пока в множестве не окажется частиц.
2. Используя модель динамики первого порядка , спрогнозировать новые значения частиц (2.11):

(2.11)

1. Для каждой частицы :
   1. Построить цветовую гистограмму согласно (2.7).
   2. Вычислить расстояние Бхаттачаария до эталонной гистограммы . (2.8).
   3. Вычислить вес согласно функции правдоподобия (2.9).
2. Нормализовать вычисленные веса частиц (2.12):

(2.12)

1. Вычислить оценку текущего состояния как среднюю частицу полученного набора :

(2.13)

Блок-схема описанного алгоритма приведена на Рис. 2.1.

|  |
| --- |
| Condensation2.png |
| Рис. 2.1. Блок-схема алгоритма воспроизведения условной плотности |

## Адаптация стохастической компоненты модели динамики

Описанный выше алгоритм отслеживания сталкивается с рядом проблем. Во-первых, модель динамики первого порядка не позволяет напрямую учитывать ускорение движущегося объекта, только с помощью вводимой погрешности. При этом динамическая составляющая ошибки с предыдущих кадров оказывает влияние на статическую часть вектора состояния на текущем кадре, уменьшая тем самым точность определения положения и размера объекта. Чтобы вклад динамической части стохастической компоненты модели динамики был достаточным, чтобы учитывать изменения скорости объекта, но при этом не оказывал значительного влияния на ее статическую часть, значения стандартных отклонений должны как минимум на один порядок превышать значения стандартных отклонений  [4].

Во-вторых, так как фильтр распространяет частицы в ограниченной области пространства состояний, могут возникнуть проблемы с отслеживанием объекта, движущегося по направлению к камере или от нее, поскольку в этом случае будут наблюдаться значительные изменения размера занимаемой объектом области кадра. Чтобы учесть эти изменения, необходимо, чтобы им были сопоставлены аналогичные изменения той части вектора ошибки, которая влияет на размер объекта. Также для повышения точности результатов отслеживания необходимо уменьшать область распространения частиц, где это возможно.

Для борьбы с выделенными проблемами предлагается использовать адаптацию стохастической компоненты уравнения движения , вычисляя ее на каждом кадре в зависимости от полученных результатов отслеживания, текущего размера объекта и заданного вектора начальных стандартных отклонений , зависящего от конкретной задачи отслеживания. Адаптация к размеру объекта может быть задана линейным соотношением, а для установления зависимости между стандартными отклонениями стохастической компоненты модели динамики и результатами работы алгоритма на кадре может быть использована сигмоидальная функция  [4]:

(2.14)

где — расстояние между эталонной цветовой гистограммой и гистограммой , построенной для вычисленной на кадре области, занимаемой объектом.

Основная идея адаптации стохастической компоненты заключается в приближении модели динамики объекта к модели случайных блужданий при снижении качества отслеживания. Другими словами, чем дальше находится рассчитанная на кадре гистограмма от эталонной гистограммы, тем больше вероятность, что объект потерян трекером. Следовательно, уже нельзя полагаться на полученные с предыдущих кадров сведения о характере движения объекта, и фильтр должен начать распространять частицы в более широкой области кадра, чтобы как можно быстрее восстановить свою работоспособность. Для этого необходимо как можно сильнее снизить влияние динамической компоненты вектора состояния объекта, увеличив при этом вклад его статической части. В противном случае, когда расстояние между вычисленной на кадре цветовой гистограммой и эталонной гистограммой мало, то с большой долей вероятности трекер находится на объекте, поэтому он может опираться на историю отслеживания для определения следующего возможного состояния объекта. В этом случае можно повысить вклад динамической составляющей вектора состояния объекта.

Таким образом, вектор стандартных отклонений может быть получен исходя из соотношений (2.15):

(2.15)

Для ускорения перехода к модели случайных блужданий при потере объекта дополнительно можно ввести искусственное взвешивание динамической компоненты вектора состояния объекта (2.16):

(2.16)

Выражение (2.16) оказывает влияние на динамическую часть вектора состояния объекта только в том случае, когда значение сигмоидальной функции близко к 1, что сигнализирует об увеличении расстояния до целевой гистограммы. В противном случае значение введенного коэффициента стремится к 1 и не влияет на значение вектора **.**

## Выводы

В данном разделе был рассмотрен в общем виде алгоритм воспроизведения условной плотности, реализующий основные принципы фильтра частиц. Были выбраны способы представления объекта на кадре и описания его внешних признаков, а также модель динамики объекта, описан процесс измерения состояния объекта и задана функция правдоподобия. Была представлена конкретизация алгоритма отслеживания с учетом описанных параметров. Также была предложена модификация разобранного алгоритма, позволяющая адаптировать вектор случайной ошибки модели динамики к получаемым на каждом кадре результатам.

# Технологический раздел

Для разработки программного обеспечения, реализующего предложенный метод отслеживания, был выбран язык C++ и графическая библиотека OpenCV.

На рис. 3.1. представлена структура разработанного программного обеспечения в виде диаграммы компонентов.

|  |
| --- |
| components.png |
| Рис. 3.1. Диаграмма компонентов разработанного программного обеспечения |

# Исследовательский раздел

## Формирование набора данных

Для проведения исследования работоспособности предложенного метода отслеживания был использован набор размеченных видепоследовательностей BoBoT (Bonn Benchmark on Tracking) [7], находящийся в открытом доступе по адресу <http://www.iai.uni-bonn.de/~kleind/tracking/> (ссылка действительна на дату 30.05.2014). Данный набор состоит из размеченных видеозаписей, для каждого кадра указаны координаты левого верхнего угла описывающего прямоугольника, его ширина и высота.

Имеющиеся в наборе видеозаписи можно разделить на две группы. На видеопоследовательностях первой группы представлен неподвижный объект, снятый перемещающейся камерой. Эти видеозаписи позволяют проверить работоспособность метода отслеживания в следующих условиях:

* изменение освещения сцены;
* резкие и хаотичные перемещения камеры;
* частичное перекрытие объекта элементом сцены.

Видеозаписи второй группы позволяют применить предложенный метод к отслеживанию перемещающегося объекта и провести исследования его поведения в следующих ситуациях:

* движение на изменяющемся неоднородном фоне;
* движение с ускорениями в различных направлениях;
* полные перекрытия объекта;
* наличие в сцене похожих объектов (дистракторов, англ. distractor).

## Метрики, используемые для оценки работы метода

С условием наличия размеченных кадров, для оценки точности определения занимаемой объектом области кадра можно использовать индекс  , предложенный в работе [10]:

(4.1)

где и — описывающие прямоугольники для размеченного и полученного на кадре состояний объекта, соответственно, — площадь прямоугольника . Таким образом, данная метрика представляет собой отношение площади пересечения описывающих прямоугольников для реальных и вычисленных состояний объекта к максимальной из их площадей. Для оценки качества отслеживания по всей видеозаписи берется среднее значение величины :

(4.2)

где — количество кадров в видеозаписи. Считается, что объект потерян на кадре , если .

Величина  включает в себя в том числе и те кадры, на которых объект был потерян трекером, поэтому она не может дать адекватной оценки точности определения области, занимаемой объектом на кадре, при наличии полных перекрытий отслеживаемого объекта другими элементами сцены. Поэтому вводится дополнительная оценка  (4.3), вычисляющая среднее значение величины только для тех кадров, на которых объект был успешно определен.

(4.3)

где , —количество кадров, на которых объект был захвачен трекером.

Для определения способности трекера восстанавливаться после потери объекта вычисляется среднее время восстановления . Эта величина определяется как усредненное по всей видеопоследовательности количество кадров, прошедших с момента потери объекта до его обнаружения трекером [4].

## Результаты экспериментов

Результаты проведенных экспериментов представлены как в графическом, так и в аналитическом виде. Графики отражают изменение точности определения области объекта в ходе видеопоследовательности.

На Рис. 4.1 представлены результаты, полученные для первой группы видеозаписей, на которых неподвижный объект снимается перемещающейся камерой.

|  |
| --- |
| nonmovingObjectQualityPlots.jpg |
| Рис. 4.1. Графичекое представление индекса качества для неподвижных объектов, снятых перемещающихся камерй |

Из представленных графиков видно, что с проблемой изменения освещения сцены справились оба алгоритма. Однако в ситуации с частичным перекрытием объекта видно небольшое преимущество предложенного метода с адаптацией стохастической компоненты модели динамики, поскольку простой алгоритм воспроизведения условной плотности в момент частичного перекрытия был на грани потери объекта, как это видно из графика. С ситуацией быстрых и хаотичных перемещений камеры стандартный вариант алгоритма не справился вообще, потеряв объект практически в начале отслеживания, в то время как график индекса точности для модифицированного алгоритма показывает, что к середине видеозаписи наблюдается некоторое падение точности определения области объекта, однако ни на одном кадре объект потерян не был. Численные значения используемых метрик для обоих алгоритмов приведены в табл. 4.1. Данные значения были получены как усредненные результаты для 20 запусков видеозаписи.

*Таблица 4.1*

Количественные оценки работоспособности алгоритмов для неподвижных объектов, снятых перемещающейся камерой.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Условия отслеживания** |  | **Без адаптации** | | | | **С адаптацией** | | | |
|  |  |  |  |  |  |  |  |
| **Резкие перемещения камеры** | 404 | 0,09 | 41,65 | 0,7 | 233,99 | 0,59 | 345,4 | 0,69 | 39,37 |
| **Изменение освещения сцены** | 412 | 0,69 | 385,85 | 0,72 | 118,5 | 0,75 | 412 | 0,75 | — |
| **Частичное перекрытие объекта** | 305 | 0,73 | 301,65 | 0,73 | 6,94 | 0,81 | 305 | 0,81 | — |

На Рис. 4.2 представлены аналогичные графики, полученные в результате применения обоих алгоритмов отслеживания к отслеживанию движущегося объекта. Из анализа графиков следует, что оба алгоритма успешно справились с задачей ускорения объекта и изменения направления его движения. Резкие точечные спады показателя качества (примерно 200 кадр) говорят о том, что в этот момент наблюдалось внезапное ускорение объекта, которое оба алгоритма не успели обработать за один кадр, однако в обоих случаях восстановление заняло всего 1 кадр.

Похожие результаты показали оба алгоритма и в ситуации отслеживания объекта, перемещающегося на неоднородном меняющемся фоне. Как следует из таблицы 4.2, их средние численные показатели также практически одинаковы. А вот с повторяющимися перекрытиями объекта лучше справился модифицированный алгоритм. Из графика хорошо видны области перекрытия — те участки, на которых значение величины резко падает, а затем опять восстанавливается. Результаты эксперимента показывают, что предложенный метод отслеживания с адаптацией стохастической компоненты уравнения динамики успешно справился со всеми перекрытиями объекта в ходе видеозаписи, в то время как стандартный алгоритм воспроизведения условной плотности перестал показывать адекватные результаты после первого же перекрытия.

Для ситуации с наличием в сцене дистракторов оба метода показали схожие результаты. На графиках видны два резких спада в районе 200 кадра. Они соответствуют ситуациям, когда объект-помеха двигался в отличном от целевого объекта направлении. Видно, что такую ситуацию оба метода обработали без труда. Достаточно длинная последовательность кадров с потерей объекта ближе к концу видеозаписи соответствует ситуации, когда объект-помеха двигался в том же направлении, что и целевой объект, находясь при этом рядом с ним.

|  |
| --- |
| movingObjectQualityPlots.jpg |
| Рис. 4.2. Графичекое представление индекса качества для движущихся объектов |

*Таблица 4.2*

Количественные оценки работоспособности алгоритмов для отслеживания перемещающихся объектов

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Условия отслежвания** |  | **Без адаптации** | | | | **С адаптацией** | | | |
|  |  |  |  |  |  |  |  |
| **Ускорение, изменение траектории движения** | 602 | 0,78 | 601,15 | 0,78 | 1 | 0,81 | 602 | 0,78 | — |
| **Изменяющийся неоднородный фон** | 629 | 0,65 | 229 | 0,65 | — | 0,66 | 629 | 0,66 | — |
| **Полные перекрытия объекта** | 453 | 0,3 | 170,45 | 0,57 | 59,79 | 0,45 | 335,9 | 0,55 | 10,06 |
| **Наличие похожих объектов в сцене** | 1017 | 0,52 | 764,25 | 0,65 | 49,63 | 0,54 | 767,65 | 0,66 | 47,93 |

## Исследование времени обнаружения объекта с момента инициализации

В начале отслеживания для первого кадра видеозаписи необходимо задать исходный набор частиц. Если положение целевого объекта заранее известно, или задается вручную оператором, то исходный набор частиц можно сгруппировать вокруг точки, принадлежащей к объекту. Для всех полученных ранее результатов использовался именно такой способ инициализации.

Если сведений о первоначальном положении объекта нет, набор частиц можно равномерно распределить по области кадра. В этом случае с точки зрения оценки качества метода отслеживания интересен вопрос, сколько кадров понадобится трекеру для обнаружения объекта. Также можно проверить, как поведут себя оба алгоритма, если изначально положение объекта задано неверно, то есть частицы сгруппированы вокруг точки, заведомо не принадлежащей к объекту.

Таблица 4.3 показывает среднее время обнаружения объекта стандартным алгоритмом воспроизведения условной плотности и его модификацией для случаев равномерного распределения частиц по кадру и ошибочного задания объекта. Приведенные результаты были получены для 100 запусков алгоритма на одной и той же видеопоследовательности.

*Таблица 4.3*

Время обнаружения объекта в зависимости от способа инициализации

|  |  |  |
| --- | --- | --- |
| **Исходное положение частиц** | **Время обнаружения объекта (в кадрах)** | |
| **Без адаптации** | **С адаптацией** |
| **Равномерно распределены по области кадра** | 39,51 ± 22,3 | 2,25±0,4 |
| **Сгруппированы вокруг точки, не относящейся к объекту** | 159,53±60,7 | 18,99±1,7 |

Эксперимент показал, что стандартному алгоритму требуется в десятки раз больше времени, чтобы обнаружить объект, если он не задан явным образом. Алгоритм с адаптацией стохастической компоненты модели динамики объекта справляется с этой задачей в среднем за 2 кадра при равномерном распределении частиц по кадру, и тратит около 20 кадров на обнаружение при условии, что частицы изначально сгруппированы вокруг точки, не имеющей отношения к объекту.

# Заключение

В ходе работы были решены все поставленные задачи:

* проведена классификация существующих подходов к визуальному отслеживанию объектов, проанализированы их достоинства и недостатки;
* сформулировано математическое описание основных принципов выбранного в качестве основы для разрабатываемого метода фильтра частиц
* разработан алгоритм отслеживания, модифицирующий стандартный алгоритм воспроизведения условной плотности, реализующий основные принципы фильтра частиц, путем введения адаптации стохастической компоненты модели динамики объекта к получаемым результатам;
* разработано программное обеспечение, реализующее предложенный метод;
* проведено исследование работоспособности разработанного метода в различных условиях отслеживания.

Результаты исследования показали, что предложенная модификация позволяет сделать стандартный алгоритм более устойчивым к перекрытиям объекта, снизить время восстановления после потери объекта, а также время, требуемое на обнаружение объекта при инициализации трекера, и повысить среднюю точность определения занимаемой объектом области на кадре видеозаписи. Основным ограничением разработанного метода является предположение, что цветовая гистограмма отслеживаемого объекта известна заранее и отлична от фона.

В качестве дальнейшего направления работы можно предложить следующие варианты повышения точности и качества отслеживания:

* использование отличного от цветовых гистограмм представления объекта, например, с помощью контуров [11], признаков Хаара, гистограмм направленных градиентов и локальных бинарных шаблонов [5];
* моделирование формы и размера объекта только с помощью частиц, исключив эти параметры из вектора состояния [3];
* обновление эталонной цветовой гистограммы объекта по ходу отслеживания [8].

# Список литературы

1. Arulampalam M. S., Maskell S., Gordon N., Clapp T. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking // IEEE Transactions on Signal Processing, 2002, vol. 50, no. 2, pp. 174-188. DOI: 10.1109/78.978374.

. Cai Y., de Freitas N., Little J. Robust visual tracking for multiple targets. // Computer Vision – ECCV, 2006, vol.3954, pp. 107-118. DOI: 10.1007/11744085\_9.

. Deardena A., Demirisa Y., Graub O. Tracking football player movement from a single moving camera using particle filters // Proceedings of the 3rd European Conference on Visual Media Production (CVMP2006), 2006, pp. 29-37.

. Del Bimbo A., Dini F. Particle filter-based visual tracking with a first order dynamic model and uncertainty adaptation // Computer Vision and Image Understanding. 2011. Vol. 115. No. 6. P. 771-786. DOI: 10.1016/j.cviu.2011.01.004.

. Grabner H., Grabner M., Bischof H. Real-time tracking via on-line boosting // Proceedings on British Machine Vision Conference (BMVC), 2006, vol. 1, pp. 47-56.

. Jaward M., Mihaylova L., Canagarajah N., Bull D. Multiple object tracking using particle filters // Aerospace Conference. IEEE, 2006, pp. 8–. DOI: 10.1109/AERO.2006.1655926.

. Klein D., Schulz D., Frintrop S., Cremers A. Adaptive real-time video-tracking for arbitrary objects // IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Taipei 2010. Pp. 772 – 777. DOI: 10.1109/IROS.2010.5650583.

. Nummiaro K., Koller-Meier E., Van Gool L. An adaptive color-based particle filter // Image and Vision Computing, 2003. vol. 21, no. 1, pp. 99–110. DOI: 10.1016/S0262-8856(02)00129-4.

. Pérez P., Hue C., Vermaak J., Gangnet M. Color-Based Probabilistic Tracking // ECCV '02 Proceedings of the 7th European Conference on Computer Vision-Part I. London 2002. Pp. 661–675.

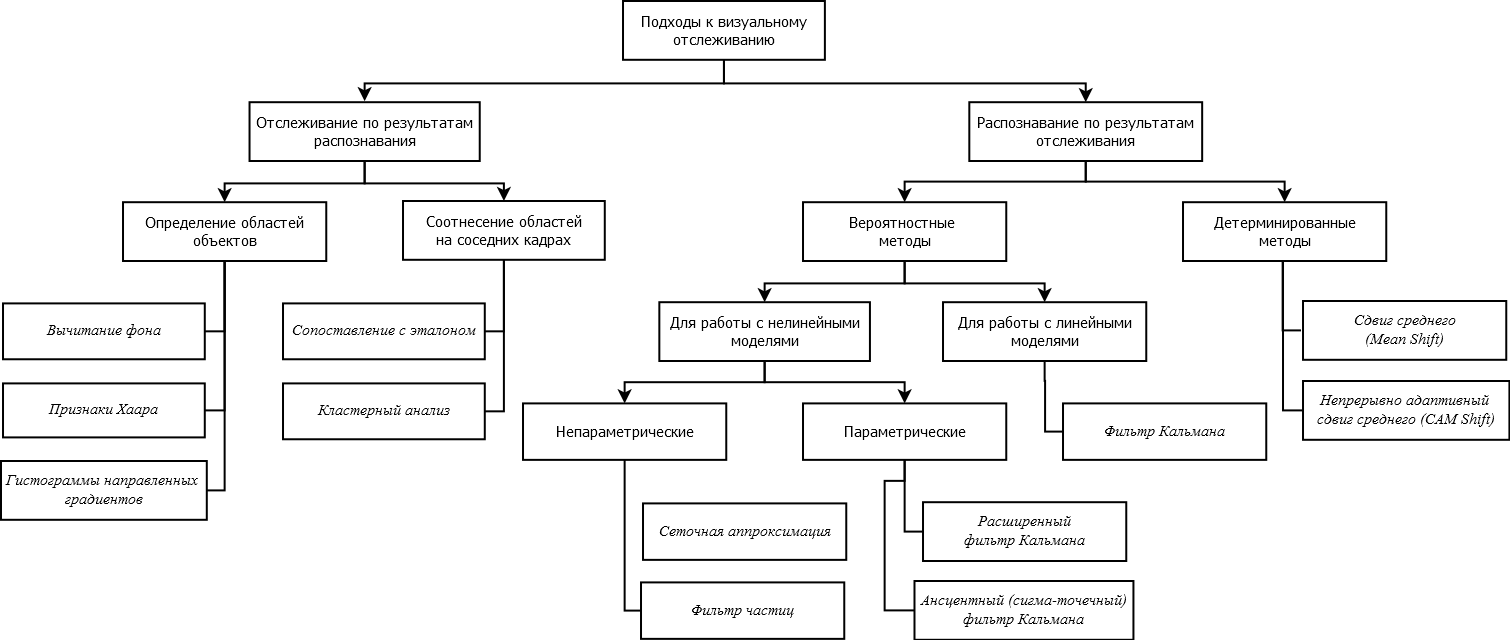
10. Phillips I. T., Chhabra A. K. Empirical performance evaluation of graphics recognition systems // IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999, vol. 21, no. 9, pp. 849-870. DOI: 10.1109/34.790427.

. Rui. Y., Chen Y. Better proposal distributions: object tracking using unscented particle filter // Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, (CVPR 2001), 2001, vol. 2, pp.II-786 – II-793. DOI: 10.1109/CVPR.2001.991045.

12. Yilmaz A., Javed O., Shah M. Object tracking: A survey // ACM Computing Surveys (CSUR), 2006, vol. 38, no. 4. DOI: 10.1145/1177352.1177355.

# Приложение А

# Классификация подходов к визуальному отслеживанию объектов



1. В названии алгоритма приведена транслитерация оригинального названия, поскольку устоявшегося термина в русском языке пока нет, а дословный перевод не имеет смысла (unscented – «без запаха»). Также данный алгоритм иногда называют сигма-точечным фильтром Кальмана (англ. Sigma-Point Kalman Filter) [↑](#footnote-ref-1)