Минобрнауки России

Федеральное государственное автономное образовательное учреждение высшего образования   
«Национальный исследовательский университет   
«Московский институт электронной техники»

Факультет микроприборов и технической кибернетики

Кафедра высшей математики №1

Димаков Владислав Сергеевич

Бакалаврская работа   
по направлению 01.03.04 «Прикладная математика»

Обнаружение и слежение за объектами в реальном времени на основе самообучающегося классификатора

Студент Димаков В.С.

Научный руководитель,

доцент, кандидат физико-математических наук Козлитин И.А.

Москва 2017

Оглавление

[Введение 4](#_Toc482366452)

[Актуальность проблемы 4](#_Toc482366453)

[Цели и задачи выпускной квалификационной работы 4](#_Toc482366454)

[Глава 1. Обнаружение и сопровождение объектов, присутствующих на кадрах видеопоследовательности. 4](#_Toc482366455)

[1.1.Цифровое растровое полутоновое изображение 4](#_Toc482366456)

[Цифровая видеозапись 5](#_Toc482366457)

[1.Виды алгоритмов обнаружения и сопровождения (кратко) : пороговые, корреляционные, вычитание фона и др. 5](#_Toc482366458)

[Вычитание фона 5](#_Toc482366459)

[2. Обнаружение и сопровождение на основе вычитания фона с учетом движения камеры 5](#_Toc482366460)

[Особые точки изображения 5](#_Toc482366461)

[Детектор Харриса 6](#_Toc482366462)

[Вычисление оптического потока методом Лукаса — Канаде 8](#_Toc482366463)

[Смещение кадра 9](#_Toc482366464)

[Модель фона 9](#_Toc482366465)

[Бинарное изображение подвижных объектов 11](#_Toc482366466)

[Сегментация изображения 13](#_Toc482366467)

[Глава 2. Распознавание образов и классификация. 13](#_Toc482366468)

[2.1. Распознавания образов. 13](#_Toc482366469)

[2.2. Виды классификаторов. 14](#_Toc482366470)

[2.3. Случайные леса. 14](#_Toc482366471)

[2.3.1. Деревья решений. 14](#_Toc482366472)

[2.3.2. Мера загрязненности вершины. 15](#_Toc482366473)

[2.3.3. Оптимальное расщепление вершин. 16](#_Toc482366474)

[2.3.4. Ансамбль деревьев решений. 17](#_Toc482366475)

[2.3.5. Алгоритм построения случайного леса. Алгоритм классификации. 18](#_Toc482366476)

[2.3.5. Оценка вероятности ошибочной классификации. 18](#_Toc482366477)

[2.4. Обучение случайного леса в режиме реального времени. 19](#_Toc482366478)

[2.4.1. On-line баггинг. 19](#_Toc482366479)

[2.4.2. Построение деревьев решений в on-line режиме. 20](#_Toc482366480)

[2.4.3. Адаптивное исключение деревьев из ансамбля. 22](#_Toc482366481)

[2.4.4. Алгоритм on-line обучения. 22](#_Toc482366482)

[Глава 3. Применение методов распознавания образов к задаче сопровождения. 23](#_Toc482366483)

[3.1. Задача долгосрочного сопровождения объектов. 23](#_Toc482366484)

[3.2. Предлагаемый метод решения задачи. 24](#_Toc482366485)

[3.3. Алгоритм работы детектора. 24](#_Toc482366486)

[3.4. Признаки цифрового изображения. 25](#_Toc482366487)

[Глава 4. Результаты работы программы 27](#_Toc482366488)

[Заключение. Перспективы дальнейшей работы. 27](#_Toc482366489)

# Введение.

## Актуальность проблемы.

Применение автосопровождения в жизни(видеонаблюдение, военное дело и т.п)

В настоящее время от систем видеонаблюдения требуется не только предоставление возможности воспроизведения и записи видеопотока с камеры, но и возможности решения в автоматическом режиме множество задач без участия человека, начиная от простого детектирования движения в области наблюдения, заканчивая высокоточным подсчётом проехавших машин или прошедших людей.

Большинство задач, решаемых системами видеонаблюдения, направлены на получение различных данных об объектах в области наблюдения, поэтому для сбора и последующей обработки информации наиболее важным вопросом является определение положений требуемых объектов на каждом кадре видеопотока.

## Цели и задачи выпускной квалификационной работы.

Целью работы является исследование и анализ эффективности методов обнаружения и слежения за объектами в реальном времени на основе данных получаемых из видеопотока.

# Глава 1. Обнаружение и слежение за объектами видеопоследовательности.

## 1.1. Видеопоследовательность.

Одной из важнейших компонент системы автоматического обнаружения и слежения является датчик изображений, поскольку именно он служит главным, а зачастую и единственным поставщиком информации, необходимой для решения задачи автоматического обнаружения и слежения за объектами. Объем информации, генерируемой датчиком, определяется количеством элементов светочувствительной матрицы датчика, количеством спектральных каналов наблюдения и частотой формирования кадров.

Наибольшее распространения в системах слежения получили одноканальные датчики видимого диапазона. Как правило, эти датчики формируют на выходе цифровой видеопоток, который представляет собой последовательность цифровых изображений.

Цифровым изображением называется массив данных, полученный путем аналого-цифрового преобразования сигнала, поступающего от датчика изображений. Существуют три основных способа представления цифровых изображений: растровая графика, векторная графика и фрактальная графика. Растровое изображение – это изображение, представляющее собой двумерную матрицу , каждый элемент которой характеризует цвет соответствующего пикселя. Пикселем называется наименьший логический элемент цифрового растрового изображения. Разрешением цифрового растрового изображения называется величина, определяющая количество пикселей изображения на единицу площади. Каждый пиксель растрового изображения может кодироваться различным количеством бит, что определяет количество возможных значений яркости.

Изображение оптического поля, получаемое с помощью датчика изображений, подвергается дискретизации по пространственным координатам и , по времени – с частотой поступления кадров. Таким образом, на устройство анализа изображений поступает последовательность матриц чисел , называемая видеопоследовательностью, каждый элемент которой представляет собой измерение яркости в соответствующей точке на каждом кадре. Здесь и – индексы изображения по пространственным осям координат, – номер кадра.

В данной работе под видеопоследовательностью понимается последовательность цифровых растровых одноканальных изображений с разрешением пикселей и числом бит на пиксель равным

**1.2. Задача обнаружения и слежения за объектами видеопоследовательности.**

Одной из задач, решаемых системами обнаружения и слежения за объектами, является задача обнаружения движущихся или появляющихся на изображении объектов. После решения задачи первичного обнаружения объекта (автоматического или с помощью оператора), как правило, должна решаться задача отслеживания траектории движения объекта.

Отслеживание траектории может производиться двумя основными методами. Первый метод, называемый слежением, предполагает оценивание координат объекта на каждом кадре наблюдаемой видеопоследовательности без решения задачи управления перемещением датчика изображения. Последовательность координат при этом формирует траекторию движения объекта. Методы, решающие задачу сопровождения также называются методами трекинга, а устройства, способные решать эту задачу – трекерами. Второй метод, не рассматриваемый в данной работе, называется сопровождением и предусматривает наряду с определением координат объекта на каждом кадре еще и решения задачи управления перемещением датчика изображений. Цель такого управления, чаще всего, состоит в удержании объекта сопровождения в центре поля зрения видеокамеры или в исключении потери визуального контакта с объектом вследствие движения датчика изображений и объекта в пространстве.

Решение задачи обнаружения и слежения осуществляется в несколько этапов. Целью начального этапа является выделение принадлежащих объектам точек текущего кадра, то есть разделение изображения на сегменты, представляющие его смысловые области. Данный процесс называется сегментацией и может осуществляться различными методами. На последующих этапах, с учетом уже имеющейся информации, решаются задачи измерения координат и оценки параметров объектов. Последним этапом является слежение, которое состоит в оценке траектории движения объектов во времени, основываясь на анализе поведения сегментов во времени. При этом для оценки траектории используются тракторные признаки, такие как близость и соотношения сторон сегментов, соответствующих объектам на соседних кадрах видеопоследовательности.

**1.3. Сегментация изображений.**

Сегментация изображения – разделение изображения на сегменты, то есть на ряд составляющих его смысловых областей. При этом предполагается, что точки каждой из смысловых областей обладают некоторой степенью общности, которая позволяет рассматривать их как одну область. Результат сегментации принято представляться в виде бинарного изображения , имеющего столько же элементов, сколько их содержит исходное изображение. В ходе сегментации всем точкам бинарного изображения, принадлежащим одному классу, присваивается одно и то же значение. Значение этих точек выбирается из множества , где значению 1 соответствуют точки, принадлежащие областям, в которых предположительно находится объект, а значению 0 – точи, соответствующие областям, предположительно содержащим фон.

Существует немало подходов к решению задачи сегментации, обладающих своими достоинствами и недостатками. Большинство из них основаны на анализе признаков точек изображения. Признаки точек – это скалярные или векторные величины, вычисляемые с помощью яркостного поля изображения, выбор которых решающим образом влияет на правильность сегментации.

**1.3.1.Метод пороговой сегментации.**

Наиболее простым методом сегментации является пороговая обработка. Результат пороговой обработки изображения может быть описан соотношением:

где – результат сегментации, – значение порога, которое в общем случае может быть различным в каждой точке изображения.

Существуют различные подходы к определению порогов: по гистограммам яркости, с использованием моментов изображения и т.д. Основная идея данного метода заключается в предположении, что изображение содержит известное число однородных по яркости классов точек, то есть таких классов, распределения вероятностей которых унимодальны. Кроме того, считается, что граничные участки между замкнутыми областями занимают сравнительно небольшую площадь изображения. Поэтому на гистограмме им должны соответствовать межмодовые впадины, в которых устанавливаются пороги сегментации.

Использование методов пороговой обработки обычно ограничено ситуациями, при которых выполняется предположение об унимодальности распределения яркостей объектов и фона, при этом гистограмма объекта и фона должны слабо перекрываться. Для успешной пороговой сегментации эти требования должны быть выполнены, если не для всего изображения, то хотя бы для областей небольшого размера, используемых при вычислении локальных порогов. Указанные ограничения позволяют использовать методы пороговой обработки лишь в простейших случаях.

**1.3.2.Метод сегментации на основе выделения границ.**

Другим популярным методом сегментации является метод, основанный на выделении границ (контуров). Этот метод основывается на предположении о существовании резкого перепада яркости между областями изображения, соответствующими различным смысловым участкам. Для выделения границ могут быть использованы различные подходы, основным из которых является пространственное дифференцирование. Данный подход основан на предположении о том, что граничные точки имеют большую величину градиента функции :

Для того, чтобы выделить границы изображения, вначале вычисляют модуль градиента изображения

а полученное градиентное изображение подвергают пороговой обработке. После этого выделяются сегменты и оценивается их размер по осям, рассчитываются коэффициенты, характеризующие форму и другие признаки, которые могут быть использованы для принятия решения о присутствии искомого объекта на изображении и оценки его координат.

Так как растровое изображение является дискретным, при его обработке для вычисления градиента часто используют дискретные оценки производных яркости изображения, вычисленные по двум взаимно перпендикулярным направлениям.

Применение методов выделения границ далеко не всегда сразу дает хороший результат, вследствие чего требуется дополнительная обработка контурного изображения, которая включает прослеживание и уточнение контуров, восстановление разрывов в контурах, удаление одиночных точек и т.д.

К сожалению, при работе с реальными изображениями предположение о наличии резких границ между смысловыми областями изображения часто не выполняется. Кроме того, смысловые области нередко содержат внутри себя резкие перепады яркости. Эти обстоятельства ведут к низкой эффективности применения методов выделения границ при обработке таких изображений. Следует заметить, что методы выделения границ не очень хорошо зарекомендовали себя при работе на сильно зашумленном изображении, что является следствием использования операций дискретного дифференцирования.

**1.3.3. Метод сегментации на основе вычитания изображения фона.**

## Данный метод основан на построении изображения фона с последующим вычитанием его из изображения текущего кадра. Изображение фона строится на основе модели фона, состоящей из двух частей: среднего фонового изображения и изображения средних абсолютных отклонений .

Среднее фоновое изображение формируется путем осреднения последних кадров видеопоследовательности. При этом обновление среднего фонового изображения новыми кадрами происходит только теми участками, где нет движущихся объектов. Пусть – среднее фоновое изображение, сформированное к кадру ; – оператор позволяющий получить изображение, без движущихся объектов:

Тогда среднее фоновое изображение для кадра вычисляется по формуле:

Изображение средних абсолютных отклонений является характеристикой средней изменчивости фона и вычисляется путем осреднения последних абсолютных отклонений кадров от фоновых изображений . Изображение средних абсолютных отклонений так же как и среднее фоновое изображение обновляется только теми участками, где нет движущихся объектов. Пусть – изображение среднего абсолютного отклонения, сформированное к кадру . Тогда изображение средних абсолютных отклонений для кадра вычисляется по следующей формуле:

Для определения точек, принадлежащих движущемуся объекту, используется предположение о том, что для движущихся точек абсолютное отклонение их значения на текущем кадре от значения на среднем фоновом изображении превышает их значение на изображении средних абсолютных отклонений. То есть движущимися считаются такие точки , для который , где .

Бинарное изображение на кадре вычисляется на основе полученной модели фона следующим образом:

где . Таким образом, бинарное изображение принимает значение, равное единице, на тех участках, для которых абсолютное отклонение текущего кадра от среднего фонового изображения превышает значение изображения среднего абсолютного отклонения, что характерно для подвижных объектов.

**1.4. Вычисление пространственных параметров объектов.**

Выше были рассмотрены алгоритмы сегментации изображений, результатом работы которых являлись бинарные изображения. Если сегментация изображения выполнена безошибочно, то бинарное изображение содержит исчерпывающую информацию, позволяющую вычислить практически любые пространственные параметры наблюдаемых объектов.

Алгоритм вычисления пространственных параметров объектов состоит из двух частей: выделение объектов на основе бинарного изображения и вычисление пространственных параметров этих объектов.

Выделение объектов на основе бинарного изображения происходит следующим образом. Среди множества точек бинарного изображения находятся связные области, то есть области представляющие собой совокупности связных элементов бинарного изображения. Связными называются такие два соседних элемента в строке или столбце матрицы, соответствующей изображению, которые имеют одинаковое значение яркости. После чего связные области окружаются прямоугольной рамкой минимально возможного размера так, чтобы все элементы связной области попали в нее. Таким образом, рамка окружает объект исходного изображения, представленный на бинарном изображении связной областью.

Вычисление пространственных параметров объектов происходит на основе полученной рамки вокруг объекта. К числу вычисляемых параметров можно отнести: размер объекта, его координаты, координаты его центра.

## 1.5. Обнаружение и слежение на основе вычитания фона с учетом движения камеры

Рассмотрим основные причины возникновения геометрических преобразований в процессе наблюдения. Для этого обратимся к модели перспективной проекции, описывающей формирование изображения на светочувствительной матрице датчика. Очевидно, что при изменении координат точки наблюдения меняются и координаты ее образа на матрице датчика . Изменение координат может быть вызвано следующими обстоятельствами: перемещение точки в пространстве, перемещение центра системы координат в пространстве, изменение ориентации осей координат. Таким образом, основные причины возникновения геометрических преобразований связаны с перемещениями объекта наблюдения, перемещениями видеодатчика и изменением ориентации видеодатчика.

Еще одной причиной возникновения геометрических преобразований является непрямолинейность распространения света, которая возникает из-за присутствия в атмосфере различных видов неоднородностей, например перепадов температуры.

Для решения задачи выделения подвижных объектов при движущемся датчике очень важно как можно точнее оценить параметры геометрических преобразований, вызванных движением датчика.

Основные классы геометрических преобразований:

При афинном преобразовании прямая линия переходит в прямую линию. Параллельность прямых также не нарушается. Афинные:

Смещение изображения . А, Б – параметры смещения по горизонтали и вертикали. Смещение обычно возникают при изменении направления оптической оси датчика. При этом необходимо, чтобы угловая ширина области обзора датчика была небольшой, иначе будут возникать более сложные преобразования. Другой источник сдвигов – движение датчика в плоскости, параллельной плоскости наблюдаемой сцены, при неизменном направлении оптической оси.

Изменение масштаба , где лямбда, мю - коэффициент изменения масштаба

Поворот наблюдаемого объекта на угол фи

Аффинные преобразования не являются адекватной моделью, если плоскости входных изображений и поля зрения датчика не параллельны. При наблюдении плоской поверхности из различных точек пространства возникают проективные преобразования. При проективном преобразовании прямые линии сохраняются, но их параллельность нарушается.

Методы оценки параметров геометрических преобразований: корреляционный метод, спектральный метод оценки параметров геометрических преобразований, метод оценки параметров геометрических преобразований на основе выбора опорных элементов.

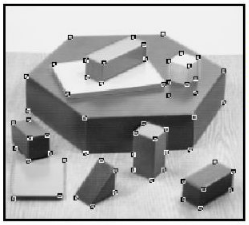
Корреляционные и спектральные методы оценки опираются на использование яркостной информации, заложенной в изображение. Альтернативный подход состоит в поиске опорных объектов на каждом из обрабатываемых изображений, определении их координат, а затем нахождении параметров геометрических преобразований.

## 1.5.1.Особые точки изображения

Для выделения из изображения некоторой интерпретируемой информации необходимо привязаться к локальным особенностям изображения.

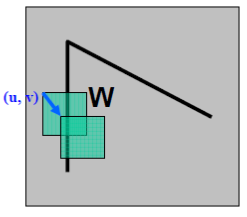
Особая точка или точечная особенность изображения – это точка изображения, окрестность которой однозначно отличима от любой другой окрестности точки , где .

Угловые особые точки (corners) – это особые точки, находящиеся на пересечении двух или более граней, которые определяют границу между различными объектами или частями одного и того же объекта. Главное свойство таких особых точек заключается в том, что в области вокруг них у градиента изображения преобладают два доминирующих направления, что делает их различимыми. Градиентом называется векторная величина, показывающая направление наискорейшего возрастания функции интенсивности изображения .

Детектор Харриса – это метод извлечения особых точек из изображения, основанный на поиске областей, похожих на угол. В нем рассматриваются производные яркости изображения по множеству направлений [3].

Для изображения рассмотрим окно (обычно его размер равен пикселей) с центром , а также его сдвиг на .



Тогда взвешенная сумма квадратов разностей между сдвинутым и исходным окном, т.е. изменение окрестности точки при сдвиге на равна:

где – весовая функция. Обычно в качестве весовой функции используется двумерный гауссиан. Так как изображение дискретно, то частные производные по оси и определяются через изменения интенсивностей соседних точек изображения. Тогда значение интенсивности сдвинутой точки можно вычислить, используя ряд Тейлора

Тогда величину можно выразить следующим образом

где *M* – автокорреляционная матрица:

Угол характеризуется большими изменениями функции по всем возможным направлениям , что эквивалентно большим по модулю собственным значениям () матрицы *M*. Поскольку напрямую считать собственные значения является трудоёмкой задачей, Харрисом и Стефеном была предложена мера отклика:

где – определитель матрицы , – след матрицы , – эмпирическая константа, ,

После вычисления меры отклика *R* производится отсечение точек по этому порогу. Далее находятся локальные максимумы функции отклика по окрестности заданного радиуса и выбираются в качестве уголковых особых точек.

## 1.5.2.Вычисление оптического потока методом Лукаса — Канаде

Оптический поток – это изображение видимого движения, представляющее собой сдвиг каждой точки между двумя изображениями. Суть вычисления оптического потока состоит в том, чтобы для каждой точки изображения найти такой сдвиг так, чтобы исходной точке соответствовала точка на втором изображении . В качестве изображений, между которыми ищется сдвиг каждой точки, будем использовать последовательность кадров цифровой видеозаписи.

Метод Лукаса-Канаде основан на предположении, что значения пикселей переходят из кадра в кадр без изменений. Таким образом, мы делаем допущение, что пиксели, относящиеся к одному и тому же объекту, могут сместиться в какую либо сторону, но их значение останется неизменным, то есть .

Используя сделанное допущение и разложение в ряд Тейлора, получаем следующее равенство

Поскольку между двумя кадрами проходит единичный интервал времени, то

Откуда получаем следующее равенство

Предположим, что соседние пиксели смещаются на одинаковое расстояние. Возьмем фрагмент изображения, пикселей, и условимся, что для каждого из пикселей и равны. Тогда вместо одного уравнения мы получим уравнений. Очевидно, что в общем случае система не имеет решения, поэтому будем искать такие и , которые минимизируют ошибку:

где – весовая функция. Обычно в качестве весовой функции используется двумерный гауссиан.

## 1.5.3.Смещение кадра

При инициализации детектора происходит создание выборки , где - координат особых точек на кадре . Используя выборку , соответствующую кадру , формируются выборки и , где , значения смещений точек на кадре относительно кадра вычисляются с помощью метода Лукаса-Канаде. Смещение кадра вычисляется, как медиана выборки . После чего кадр смещается на величину по горизонтали и на по вертикали с помощью параллельного переноса. Так как значения смещений, полученных данным образом, могут иметь дробные части, выполняется субпиксельное смещение с помощью билинейной интерполяции, используя следующую формулу

где – дробные части и ,

|  |  |
| --- | --- |
|  |  |

## 1.5.4. Принцип работы.

Модель фона состоит из двух частей – среднего фонового изображения и изображения среднего абсолютного отклонения .

Среднее фоновое изображение вычисляется путем осреднения последних кадров с учетом их взаимных смещений. При этом обновление среднего фонового изображения новыми кадрами происходит только теми участками, где нет движущихся объектов. Смещение среднего фонового изображения выполняется так, чтобы область видимости на среднем фоновом изображение максимально соответствовала области видимости на текущем кадре.

Пусть – среднее фоновое изображение, сформированное к кадру ; – смещение кадра относительно кадра ; – оператор наложения среднего фонового изображения на кадр с учетом смещения ; – оператор позволяющий получить изображение, представляющее собой кадр без движущихся объектов. Тогда среднее фоновое изображение для кадра вычисляется по следующей формуле:

Изображение среднего абсолютного отклонения является характеристикой средней изменчивости фона и вычисляется путем осреднения последних абсолютных отклонений кадров от фоновых изображений . Изображение среднего абсолютного отклонения так же как и среднее фоновое изображение обновляется только теми участками, где нет движущихся объектов, и смещается вслед за областью видимости на текущем кадре.

Пусть – изображение среднего абсолютного отклонения, сформированное к кадру . Тогда изображение среднего абсолютного отклонения для кадра вычисляется по следующей формуле:

где изображение представляет собой матрицу того же размера, что и , заполненную значениями 255 (т.е. белый цвет).

Оператор возвращает только те участки изображения, для которых на текущем кадре абсолютное отклонение от среднего фонового изображения не превышает значения изображения среднего абсолютного отклонения, умноженного на константу . То есть обновляются только те участки изображения, для которых

## 1.5.5. Глобальный порог.

Фильтрация изображений составляет основу большинства методов обработки и анализа изображений. Она не только является частью алгоритмов обнаружения и выделения объектов, но применяется и для подавления шумов, и для реставрации, восстановления, подчеркивания деталей изображений.

У вышеописанного метода вычисления бинарного изображения подвижных объектов есть один существенный недостаток. Из-за зашумленности исходного изображения некоторые пиксели, принадлежащие фону, имеют большое отклонение от среднего фонового изображения, вследствие чего такие пиксели на бинарном изображении имеют значение «1». Данный недостаток устраняется введением фильтра фоновой части изображения , работающего по следующему принципу:

где – порог фоновой части изображения, вычисляемый следующим образом:

где – гистограмма изображения ;

|  |  |
| --- | --- |
| C:\Users\Vlad\Desktop\ВУЗ\4 курс\8 семестр\Диплом\cv2\ВКР\detector\current_frame.jpg  Кадр , полученный с видеокамеры | C:\Users\Vlad\Desktop\ВУЗ\4 курс\8 семестр\Диплом\cv2\ВКР\detector\average_backgroung.jpg  Среднее фоновое изображение |
| C:\Users\Vlad\Desktop\ВУЗ\4 курс\8 семестр\Диплом\cv2\ВКР\detector\deviation_image.jpg  Масштабированное по яркости изображение среднего абсолютного отклонения | C:\Users\Vlad\Desktop\ВУЗ\4 курс\8 семестр\Диплом\cv2\ВКР\detector\current_deviation_image.jpg  Масштабированное по яркости изображение |
| C:\Users\Vlad\Desktop\ВУЗ\4 курс\8 семестр\Диплом\cv2\ВКР\detector\moving_target2.jpg  Бинарное изображение подвижных объектов | C:\Users\Vlad\Desktop\ВУЗ\4 курс\8 семестр\Диплом\cv2\ВКР\detector\moving_target.jpg  Бинарное изображение, полученное с применением фильтра фоновой части изображения |
| C:\Users\Vlad\Desktop\ВУЗ\4 курс\8 семестр\Диплом\cv2\ВКР\detector\hist1.emf | C:\Users\Vlad\Desktop\ВУЗ\4 курс\8 семестр\Диплом\cv2\ВКР\detector\hist2.emf |
| Гистограмма изображения | |

# Глава 2. Распознавание образов и классификация.

## 2.1. Распознавания образов.

Распознавание образов – это научная дисциплина, целью которой является классификация объектов по нескольким категориям или классам. Распознаваемые объекты называются образами и представляют собой вектора признаков . Признаком называется некоторое количественное измерение объекта, которое может быть выполнено в различных шкалах: числовой, порядковой или номинальной. Множество всех возможных значений векторов признаков называется пространством образов. Ответы распознавателя представляют собой вектора из некоторого пространства ответов .

Используя вероятностную постановку задачи классификации, задачу распознавания образов можно записать следующим образом. Пусть множество пар является вероятностным пространством с неизвестной вероятностной мерой . Имеется конечная обучающая выборка наблюдений , сгенерированная согласно вероятностной мере . Требуется построить алгоритм , способный классифицировать произвольный объект . Здесь – вектор, координаты которого представляют собой значения признаков, измеренных на некотором объекте; – метка класса, ; – число классов.

## 2.2. Виды классификаторов.

## 2.3. Случайные леса.

Случайный лес – статистический метод, предназначенный для решения задач классификации и регрессии, заключающийся в использовании ансамбля классификаторов, представляющих собой деревья решений. Метод сочетает в себе две основные идеи: метод баггинга и метод случайных подпространств.

### 2.3.1. Деревья решений.

Пусть – конечное непустое множество, – конечное множество, состоящее из поименованных неупорядоченных пар элементов множества . Совокупность множеств и называют неориентированным графом и обозначают . Элементы называют вершинами, а элементы – ребрами. Если – это неупорядоченная пара вершин и , то пишут , причем вершины и называют концами ребра .

Пусть – конечное множество, состоящее из поименованных упорядоченных пар элементов множества . Тогда совокупность множеств и называют ориентированным графом (орграфом) и обозначают . Элементы называют ориентированными ребрами или дугами. Если дуга есть упорядоченная пара вершин и , то пишут , причем вершину называют началом дуги , а вершину ее концом. Число дуг, исходящих из вершины орграфа , называют степенью исхода вершины , а число дуг, заходящих в вершину , называют степенью захода вершины .

Путем длины на графе из вершины в вершину называется последовательность или вершин и ребер графа, в которой . Связным называется граф, между любой парой вершин которого существует как минимум один путь. Циклом называется замкнутый путь ненулевой длины без повторяющихся ребер. Граф без циклов называется ациклическим.

Дерево – связный ациклический граф. Ориентированным деревом называется ациклический связный орграф, в котором только одна вершина имеет нулевую степень захода, а все остальные вершины имеют степень захода 1. Вершина с нулевой степенью захода называется корнем дерева, а вершины с нулевой степенью исхода называются листьями или терминальными вершинами.

Для каждого листа в ориентированном дереве существует единственный путь из корня. Максимальная длина пути из корня в лист называется высотой дерева. Если из вершины ведет дуга в вершину , то называется отцом, а – сыном.

Деревом решений называться ориентированное дерево , с каждой вершиной которого связаны:

1. Подмножество ;
2. Подвыборка , такая, что ;
3. Функция (правило) определяющая разбиение множества на непересекающихся подмножеств. Здесь – количество потомков вершины .

Заметим, что с корневой вершиной связывается все пространство образов и вся выборка . С листьями дерева не связываются никакие функции .

Пусть – вершина, являющаяся -м сыном вершины . Тогда множество и функция определяют подмножества следующим образом:

Цель построения дерева решений состоит в классификации векторов из распределения . Процесс принятия решений начинается с корневой вершины и состоит в последовательном применении правил , связанных с вершинами дерева. Результатом этого процесса является определение листа такого, что , при этом вектор относится к классу, являющемуся мажорантным в подвыборке .

### 2.3.2. Мера загрязненности вершины.

Алгоритм построения дерева решений основан на идее рекурсивного разбиения обучающей выборки на две более однородные подвыборки с помощью одного из признаков. Обычно вместо меры однородности используется противоположная по смыслу мера загрязненности.

Пусть – мера загрязненности вершины , – подвыборка, связанная с этой вершиной. Необходимо потребовать, чтобы загрязненность вершины была равна , если содержит прецеденты только одного класса и была максимальной в случае, если содержит одинаковое число прецедентов каждого класса. Одной из наиболее часто используемых является мера загрязненности вершины, формализованная в индексе Gini:

где – доля примеров класса в подвыборке . Другой популярной мерой является мера загрязненности, основанная на понятии энтропии:

Реже применяется мера загрязненности, основанная на частоте ошибочной классификации:

В работе [1] было показано, что выбор меры загрязненности не оказывает существенного влияния на точность классификации, поэтому, в силу простоты реализации, в качестве используемой в данной работе меры загрязненности был выбран индекс Gini.

### 2.3.3. Оптимальное расщепление вершин.

Правило разбиения множества , связанное с каждой вершиной дерева решений, называется расщеплением. Бинарное расщепление вершины можно рассматривать как функцию , где в случае вектор относится к левому сыну, а в случае – к правому.

Расщепление подвыборки необходимо осуществлять таким образом, чтобы максимально уменьшить загрязненность. Уменьшение загрязненности вершины для бинарных деревьев определяется как

где и – доли примеров подвыборки , соответствующие левому и правому сыновьям ( и ). Наилучшим расщеплением вершины считается разбиение, которое максимизирует величину , то есть оптимальное расщепление необходимо выполняться таким образом, чтобы .

Определим величину , используя в качестве меры загрязненности вершины индекс Gini. Пусть – размер подвыборки ; и – размеры подвыборок, соответствующих потомкам и ; и – число экземпляров класса в потомках и . Тогда оптимальным расщеплением вершины будем считать такое расщепление, что

### 2.3.4. Ансамбль деревьев решений.

Ансамбль классификаторов представляет собой множество классификаторов, чьи решения комбинируются некоторым образом для получения окончательной классификации наблюдений. Обычно синтез решений отдельных классификаторов, составляющих ансамбль, осуществляется путем их голосования. Основным достоинством ансамбля классификаторов пред отдельным классификатором является то, что при соблюдении определенных условий, он обладает значительно более высокой точностью, а также устойчив к «зашумлению» обучающей выборки. Необходимым и достаточным условием высокой точности ансамбля классификаторов является то, чтобы составляющие его отдельные классификаторы были достаточно точны и различны, то есть совершали ошибки на различных примерах. Для выполнения этого условия в случайных лесах применяются два метода: манипулирование примерами обучающей выборки и инъекция случайности в индуктивный алгоритм.

Для манипулирования примерами обучающей выборки используется метод, называемый баггингом. Он основан на формировании обучающей выборки для каждого дерева ансамбля с помощью бутстрепа и использовании невзвешенного голосования при агрегировании решений отдельных классификаторов. Бутстрепом называется случайная выборка с возвращением из исходной обучающей выборки, имеющая тот же объем, что и обучающая выборка.

Инъекция случайности в алгоритм построения случайных лесов происходит при построении каждого дерева решений на стадии расщепления вершин, где из всего множества признаков обучающей выборки случайно выбирается только их фиксированное число. Таким образом, на каждой итерации построения деревьев решений фактически используется метод случайных подпространств.

### 2.3.5. Алгоритм построения случайного леса. Алгоритм классификации.

Пусть – обучающая выборка; – число деревьев в ансамбле; – число случайно отбираемых признаков на этапе разделения вершин. Тогда алгоритм построения случайного леса может быть представлен в следующем виде:

Для выполнить:

1. Сформировать бутстреп выборку размера по исходной обучающей выборке ;
2. По бутстреп выборке построить неусеченное дерево решений с минимальным количеством наблюдений в терминальных вершинах равным , рекурсивно следуя следующему подалгоритму:
3. Из исходного набора признаков случайно выбрать признаков;
4. Из признаков выбрать признак, который обеспечивает наилучшее расщепление;
5. Расщепить выборку, соответствующую обрабатываемой вершине, на две подвыборки и построить двух сыновей вершины.

Классификация примера осуществляется следующим образом. Пусть – класс, предсказанный деревом решений для вектора , то есть . Тогда – класс наиболее часто встречающийся во множестве , то есть класс предсказанный случайным лесом для вектора .

### 2.3.5. Оценка вероятности ошибочной классификации.

Одно из достоинств случайных лесов состоит в том, что для оценки вероятности ошибочной классификации нет необходимости использовать кросс-проверку или тестовую выборку. Оценка вероятности ошибочной классификации случайного леса осуществляется методом «Out-Of-Bag» (OOB), состоящем в следующем.

Известно, что каждая бутстреп выборка не содержит примерно 37% примеров исходной обучающей выборки (поскольку выборка с возвращением, то некоторые наблюдения в нее не попадают, а некоторые попадают несколько раз). Будем классифицировать некоторый вектор , используя только те деревья случайного леса, которые строились по бутстреп выборкам, не содержащим . Частота ошибочно классифицированных векторов обучающей выборки при таком способе классификации представляет собой оценку вероятности ошибочной классификации случайного леса методом OOB. Практика применения оценки OOB показала, что в случае, если количество деревьев достаточно велико, то эта оценка обладает высокой точностью.

Обычно оценку вероятности ошибочной классификации, найденную методом OOB, обозначают как OOBE (Out-Of-Bag-Error). Также часто используют оценку OOB для отдельных деревьев ансамбля, обозначая ее . Данная оценка определяется, как частота ошибочной классификации векторов , не входящих в бутстреп выборку, на основе которой строилось дерево . При этом в классификации участвует не весь ансамбль деревьев, а только само дерево .

## 2.4. Обучение случайного леса в режиме реального времени.

Обычно случайные леса обучаются до начала их тестирования, то есть фазы обучения и тестирования разделены. Однако на практике обучающие данные часто не могут быть даны заранее, а поступают последовательно. Например, при решении задачи обнаружения и сопровождения объектов, присутствующих на кадрах видеопоследовательности, необходимо обучать классификатор и получать его прогнозы в режиме реального времени (on-line).

Так как случайный лес представляет собой ансамбль деревья решений, объединённых посредством баггинга, то для реализации on-line обучения случайного леса необходимо осуществлять как баггинг, так и построение самих деревьев решений в on-line режиме.

### 2.4.1. On-line баггинг.

Допустим, что обучающая выборка дана заранее и имеет размер . Тогда для построения каждого дерева случайного леса необходимо создать бутстреп выборку размера . Пусть случайная величина принимает значение 1, если пример попал в выборку на позицию и значение 0 в противоположном случае. Соответствующие вероятности равны:

Тогда – конечная последовательность независимых случайных величин, имеющих одинаковое распределение Бернулли с параметром . Случайная величина , являющаяся числом попаданий примера в выборку , имеет биномиальное распределение с параметрами и , то есть

Вероятность того, что пример попадет выборку ровно раз:

В случае последовательно поступающих данных, мы не можем определить размер обучающей выборки заранее. Будем считать, что , тогда и биномиальное распределение можно аппроксимировать распределением Пуассона с параметром , то есть

Таким образом, при баггинге в on-line режиме, последовательно поступающие данные возможно моделировать при помощи пуассоновского распределения. Для этого необходимо обновлять каждое дерево на каждом примере раз подряд, где – последовательность случайных величин, принимающих неотрицательные целочисленные значения, такая что для всякого выполняется условие

То есть – последовательность распределений асимптотически равных распределению .

### 2.4.2. Построение деревьев решений в on-line режиме.

Главной проблемой при построении дерева решений в on-line режиме является правильность выбора моментов разделения вершин. Правильность этого выбора крайне важна, так как надежность статистики, накапливаемой в последующих вершинах-сыновьях, напрямую зависит от правильности выбора момента разделения вершины-отца.

При on-line обучении статистика в вершине накапливается с течением времени и для обеспечения правильности выбора момента разделения, необходимо соблюдать два условия:

1. Число примеров, прошедших через вершину должно быть достаточным, чтобы в ней находилась надежная статистика;
2. Разделение вершины должно быть произведено так, чтобы достигалась цель верной классификации.

Для соблюдения этих условий введем в алгоритм разделения вершины две пороговых величины и , численные значения которых подбираются экспериментально так, чтобы минимизировать ошибку классификации . Здесь – минимальное число примеров, прошедших через вершину до ее разделения; – минимальное значение уменьшения загрязненности вершины такое, что , где – набор случайно выбранных признаков для вершины .

От дерева, построение которого производится в on-line режиме, также требуется возможность произведения классификации сразу же после разделения вершины, то есть до получения новых данных.

Пусть – вершина, удовлетворяющая условиям разделения с пороговыми величинами и . При прохождении примеров через эту вершину, были накоплены статистики и , представляющие собой статистики меток классов левых и правых возможных сыновей вершины , разделенных по признаку . Тогда при разделении вершины, из всего числа признаков выбирается случайный набор из признаков, после чего из набора выбирается признак , обеспечивающий наилучшее расщепление, и создаются вершины-сыновья и , разделенные по признаку . После этого значения статистик и передаются левому и правому сыновьям соответственно. Таким образом, новые вершины появляются с уже накопленной статистикой, что позволяет производить классификацию без необходимости получения новых данных.

### 2.4.3. Адаптивное исключение деревьев из ансамбля.

Часто при on-line обучении распределение примеров в обучающей выборке меняется на протяжении времени. Следовательно, необходимо иметь механизм, способный «отучивать» случайный лес от старой информации. Этого можно достичь, исключая из ансамбля деревья, неспособные верно классифицировать примеры с изменившимся распределением. Если ансамбль состоит из большого числа деревьев, то исключение из него одного дерева не приносит вреда точности классификации ансамбля в целом. Таким образом, непрерывное исключение деревьев, неспособных классифицировать примеры с изменившимся распределением, гарантирует адаптацию классификатора к изменяющимся условиям на протяжении времени.

Рассмотрим обучающий пример . Моделирование последовательно поступающих данных при помощи пуассоновского распределения ведет к тому, что некоторые деревья не участвуют в обучении на этом примере, то есть для этих деревьев число обновлений на примере равно нулю. Следовательно, на примере мы можем оценить как для всего леса, так и для дерева, не участвующего в обучении по этому примеру.

Очевидно, что деревья, неспособные верно классифицировать примеры, имеют большое значение . Следовательно, на основе значений можно производить случайное исключение деревьев из ансамбля, причем вероятность исключения -го дерева должна зависеть от .

### 2.4.4. Алгоритм on-line обучения.

Пусть – обучающий пример из последовательности поступающих данных; – число деревьев в ансамбле; Тогда алгоритм on-line обучения случайного леса может быть представлен в следующем виде:

Для каждого дерева выполнить:

1. Найти число обновлений дерева на примере . .
2. Если , то раз обновить дерево на примере . Для всех неразделенных вершин дерева , для которых выполнены условия: число примеров в вершине не менее и – следовать следующему подалгоритму:
3. Из множества выбрать признак , который обеспечивает наилучшее расщепление;
4. Разделить вершину по признаку , то есть построить ее сыновей и ;
5. Передать значения статистик и левому и правому сыновьям.
6. Если , то оценить .

# Глава 3. Применение методов распознавания образов к задаче слежения за объектами видеопоследовательности.

## 3.1. Задача долгосрочного слежения за объектами.

Задачей долгосрочного слежения называется такая задача, которая предполагает слежение за объектами в течение большого, возможно, бесконечного промежутка времени. При этом трекер, осуществляющий долгосрочное слежение должен корректно продолжать работу при сбое. Сбои работы трекера часто происходят в случае изменения окружающей обстановки, перекрытия сопровождаемого объекта другими объектами сцены или же в случае исчезновения сопровождаемого объекта из области наблюдения.

Базовые методы решения задачи слежения, описанные в главе №1, основаны на вычислении изменения координат объекта от кадра к кадру. Данные методы не позволяют решать задачу долгосрочного слежения непосредственно, так как не предполагают корректного поведения трекера в условиях сбоя.

Очевидно, что решение задачи долгосрочного слежения требует повторного нахождения отслеживаемого объекта в случае сбоя. С этой целью в алгоритм слежения часто интегрируют детектор, основанный на предварительно обученном классификаторе. Обучение таких детекторов производится на заранее подготовленных размеченных выборках, что не позволяет использовать их в том случае, если отслеживаемый объект не известен заранее. Другим недостатком таких методов является то, что фазы обучения и тестирования строго разделены, следовательно, изменения окружающей обстановки, не представленные в обучающей выборке, не могут стать частью модели. Решением проблем, связанных с предварительным обучением детектора, является использование классификатора, способного обучаться в режиме реального времени, описанного в главе №2.

Таким образом, целью данной работы является создание метода долгосрочного слежения за объектами, предполагающего наличие минимальной начальной информации. Под минимальной начальной информацией понимается то, что объект отслеживания не известен заранее и единственной информацией о нем являются его координаты, заданные пользователем на первом кадре видеопоследовательности.

## 3.2. Предлагаемый метод решения задачи.

В основе алгоритма долгосрочного слежения, разрабатываемого в рамках данной работы, лежит использование двух компонент: трекера, основанного на методе вычитания фона, и детектора, основанного на случайном лесе решений. Выше была описана модификация метода вычитания фона, позволяющая компенсировать движение камеры и осуществить фильтрацию нежелательных шумов, а также был представлен алгоритм построения случайного леса, обеспечивающий обучение в режиме реального времени.

Суть предлагаемого алгоритма заключается в следующем. Трекер отслеживает положение цели от кадра к кадру, при этом выбор объекта для отслеживания осуществляется пользователем на первом кадре видеопоследовательности. В случае если трекер теряет цель, детектор начинает ее поиск, и в случае его удачного завершения, сообщает трекеру новые координаты отслеживаемого объекта. Таким образом, использование двух компонент делает метод устойчивым к сбоям.

## 3.3. Алгоритм работы детектора.

Алгоритм работы детектора можно условно разделить на два этапа: обучение классификатора и поиск отслеживаемого объекта с его помощью.

Обучение классификатора осуществляется на последовательно поступающих примерах, формируемых трекером, представляющих собой последовательности вида , где – число выделенных на кадре областей, , – номер дерева в ансамбле. Последовательность можно представить следующим образом:

где – пример соответствующий области цели, – примеры, соответствующие областям, не являющимся областью отслеживаемого объекта. Алгоритм построения примеров последовательности следующий. Трекер выделяет на каждом кадре одну область – прямоугольную область , соответствующую цели. Изображение разбивается на непересекающуюся прямоугольную область произвольного размера. В областях вычисляются значения векторов признаков , при этом значение метки класса для вектора , вычисленного в области , принимается равным единице, а значение остальных меток классов – нулю.

В случае сбоя трекера, детектор осуществляет поиск отслеживаемого объекта следующим образом. На каждом кадре трекер, основанный на методе вычитания фона, осуществляет сегментацию изображения на области , где – число сегментов. Для каждого дерева ансамбля вычисляется последовательность векторов признаков , где вектор соответствует области . Классификатором принимается решение, какой из векторов признаков соответствует отслеживаемому объекту, после чего координаты области выбираются в качестве новых координат цели, и процесс слежения трекером за целью запускается заново.

## 3.4. Признаки цифрового изображения.

В качестве признаков, значения которых вычисляются в областях изображения , были выбраны признаки Хаара. Признаки Хаара – это признаки цифрового изображения, используемые в распознавании образов. Своим названием они обязаны сходством с вейвлетами Хаара и были впервые описаны в работе [2]. Признак Хаара состоит из смежных прямоугольных областей, позиционируемых на изображении. Основываясь на числе смежных областей, выделяют несколько типов признаков Хаара: двух-прямоугольные, трех-прямоугольные и четырех-прямоугольные.

|  |  |  |
| --- | --- | --- |
|  |  |  |
| (1) | (2) | (3) |

При вычислении значений смежных прямоугольных областей используется интегральное представление изображения, то есть матрица , совпадающая по размерам с исходным изображением, каждый элемент которой представляет собой сумму интенсивностей всех пикселей, находящихся левее и выше данного элемента:

Интегральную сумму произвольной прямоугольной области можно вычислить, используя интегральное представление изображения следующим образом:

|  |  |
| --- | --- |
|  |  |

Значение признаков Хаара первой группы вычисляется, как разность между интегральными суммами пикселей в двух смежных прямоугольных областях:. Для второй группы значение вычисляется, как интегральная сумма двух внешних областей, вычитаемая из суммы центральной области:. Значение третьей группы признаков вычисляется, как разность между суммами диагональных пар областей: .

Для каждого дерева ансамбля при инициализации классификатора строится шаблон, по которому впоследствии будут вычисляться значения векторов признаков . Этот шаблон представляет собой область фиксированного размера ( пикселей), на которую позиционируются признаков Хаара, причем размер признаков, их координаты и тип выбираются случайным образом, чтобы обеспечить различность строящихся деревьев решений.

Рассмотрим вектор признаков для дерева , вычисляемый в области . Координаты этого вектора представляют собой значения признаков Хаара, вычисляемых в области по масштабированному шаблону. Шаблон масштабируется так, чтоб совпадать размером с областью , что обеспечивает неизменность значений признаков при изменении размеров области. Значение признака принимается равным единице в случае, если разность интегральных сумм смежных прямоугольных областей неотрицательна, и нулю – в противоположном случае.

# Глава 4. Результаты работы программы

# Заключение. Перспективы дальнейшей работы.

[1]



[2] Viola Jones

[3] (http://www.bmva.org/bmvc/1988/avc-88-023.pdf)