

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
КАФЕДРА КОМПЬЮТЕРНЫХ ТЕХНОЛОГИЙ И СИСТЕМ

Егорова Екатерина Романовна

Магистерская диссертация

**Разработка системы анализа траекторий движения
пешеходов методами искусственного интеллекта по
данным камер видео наблюдения**

Направление 02.04.02

Фундаментальная информатика и информационные технологии
Магистерская программа: Автоматизация научных исследований

Научный руководитель,
доктор тех. наук,
профессор Дегтярев А.Б.

Санкт-Петербург

2019

Оглавление

Введение.....	3
Постановка цели и задачи	5
Обзор литературы	6
Глава 1. Данные и использованные инструменты.....	12
Глава 1.1 OpenPose.....	15
Глава 1.2. Преобразование координат	17
Глава 1.3. Линейная регрессия	21
Глава 1.4. Многослойный персептрон	23
Глава 2. Распознавание поз	25
Глава 3. Определение роста пешеходов в условиях потери данных	30
Глава 4. Перспективное преобразование координат.....	35
Заключение.....	43
Список литературы	44
Приложение	

Введение

Сегодня технологии искусственного интеллекта используются во многих областях: от фильтрации содержимого в социальных сетях до биржевых роботов. Они все чаще присутствуют в потребительских продуктах, таких как камеры смартфонов или камеры видеонаблюдения, где основной задачей компьютеров является идентификация объектов на изображениях. В решении таких задач все чаще приоритет отдается методам глубокого обучения. Глубокие нейронные сети привели к прорывам в распознавании образов на изображениях и видеозаписях. В течение многих лет, они доказали свою эффективность по сравнению с другими методами распознавания. Разработкой нейронных сетей занимались такие технологические гиганты, как Google (сеть-игрок в игру го AlphaGo), Microsoft (ряд сервисов для идентификации изображений), стартапы MSQRD, Prisma и другие.

С развитием методов определения положения предметов на изображениях и видеозаписях, появилась необходимость отслеживания положения и перемещения людей, с целью дальнейшего использования этой информации в обеспечении безопасности и в маркетинге. Умея распознавать позы и части тела людей, можно предсказывать девиантное поведение отдельно взятых личностей в толпе, анализировать поведение покупателей, отслеживать их перемещение в помещениях с целью дальнейшего логического размещения товаров на полках или рекламных проспектов. Более того, данные, полученные с помощью нейронных сетей можно отображать в виртуальном трехмерном пространстве. Такое решение может быть востребовано в среде разработки видеоигр, в частности, в симуляциях боев и предсказаниях поведения пользователей-игроков. Методы глубокого обучения предоставляют большие возможности для решения такого рода задач.

В данной работе предлагалось, используя методы глубокого обучения научиться распознавать части тела пешеходов на кадрах видео, взятых с камер наблюдения и научиться получать траектории перемещения пешеходов.

Постановка цели и задачи

Целью данной работы является разработка системы анализа траекторий движения пешеходов методами искусственного интеллекта по данным камер видеонаблюдения, способной работать в режиме реального времени.

Задачи:

1. Накопление тестовых данных: видео, записанное статичной камерой видеонаблюдения в торговом центре;
2. Создание системы, способной распознавать пешеходов на видео в режиме реального времени;
3. Определение координат ног (x ; y) в двумерном пространстве;
4. Выполнение проективного преобразование координат для отслеживания трекинга пешеходов и перенесения данных из двумерного в трехмерное пространство;
5. Создание универсального клиент-серверного приложения, возвращающего изображение трекинга и координаты пешеходов для отображения этих данных в трехмерном пространстве.

Актуальность данной работы обусловлена заинтересованностью служб безопасности, маркетологов, работодателей в создании системы, способной отслеживать траекторию перемещения людей в различных пространствах: на улице, в магазинах и торговых центрах, а также в школах и университетах.

Обзор литературы

Одним из самых распространенных методов решения задачи распознавания образов, являются методы глубокого обучения или нейронные сети. Существует несколько исследовательских работ, описывающих методику распознавания людей и частей тел как на изображениях, так и видео.

Одной из самых интересных стала работа, написанная в 2014 году, «Articulated pose estimation by a graphical model with image dependent pairwise relations» [8], подробно описывающая метод оценки человеческой позы по одному двумерному статическому изображению, основанный на графической модели, описанной парными соотношениями. Описывается графическая модель позы человека, которая использует тот факт, что локальные измерения изображения могут использоваться как для обнаружения деталей (или соединений), так и для прогнозирования пространственных отношений между ними (зависимые от изображения парные отношения). В работе используются глубокие сверточные нейронные сети, с помощью которых определяются условные вероятности присутствия частей и их пространственных отношений в участках изображения. Модель сочетает в себе репрезентативную гибкость графических моделей и эффективность глубоких нейронных сетей. По словам авторов, данный метод значительно превосходит современные методы значительно улучшил распознавание частей тела и показал лучший результат, опубликованный ранее, в среднем на 5,8% на наборах данных LSP и FLIC.

Специалисты Технического университета Дармштадта в 2009 году в своем исследовании «People-Tracking-by-Detection and People-Detection-by-Tracking» [5] рассматривают проблему обнаружения и отслеживания людей в условиях сложных сцен реального мира, в которых обычно участвуют несколько человек, сложные окклюзии, загроможденные или даже движущиеся фоны. Исследователями было установлено, что детекторы действительно способны находить пешеходов даже в сложных уличных сценах (многие люди частично, или также полностью закрыты в течение

длительного периода времени), но ложные срабатывания по-прежнему случаются довольно часто. Выявление отдельных лиц также остается сложной задачей. Методы отслеживания позволяют находить конкретного человека в последовательностях изображений, однако это сильно осложняется сценариями из реальной жизни, такими как переполненные улицы.

В своем исследовании, специалисты объединяют преимущества обнаружения и отслеживания в одной структуре. Приблизительная артикуляция каждого человека обнаруживается в каждом кадре на основе отдельно взятых локальных признаков, которые моделируют внешний вид отдельных частей тела. Предварительные знания о возможных сочленениях и временной когерентности в цикле ходьбы моделируются с использованием иерархической модели скрытых переменных гауссовского процесса. В работе показано, как сочетание этих результатов улучшает гипотезы о положении и артикуляции каждого человека в нескольких последующих кадрах. Представлены экспериментальные результаты, которые демонстрируют то, как происходит обнаружение и отслеживание нескольких людей в беспорядочных сценах с повторяющимися окклюзиями.

В 2017 году на базе Университета Карнеги-Меллона была разработана система распознавания частей тела «Part Affinity Fields». Модель подробно описана в работе «Realtime multi-person 2D pose estimation using part affinity fields» [7], где и представлен подход для эффективного обнаружения двумерно пространственных поз нескольких людей на изображении. В статье рассмотрен алгоритм для обнаружения двумерных поз нескольких людей на изображениях. Представлено явное не параметрическое представление ассоциации ключевых точек, которые кодируют положение и ориентацию конечностей человека. Описана разработанная архитектура для совместного обучения обнаружения деталей и ассоциации деталей. Описано действие жадного алгоритма синтаксического анализа, достаточного для качественного анализа поз тела даже при увеличении числа людей на изображении.

Еще один метод распознавания описан в совместной работе исследователей из Германии, Швейцарии и Австралии, которая называется «Human pose estimation using body parts dependent joint regressors» [10] в 2013 году. В этой работе рассматривается проблема оценки двумерной позы человека по статичным изображениям. Методы, основанные на специально обученных деформируемых деталях, организованных в виде древовидной модели, показали себя очень успешными в решении этой задачи. В рамках такой графической структуры решается проблема получения хороших шаблонов деталей, предлагая новые нелинейные регрессоры соединений. В частности, используются двухслойные случайные леса в качестве совместных регрессоров. Первый уровень действует как различающий, независимый классификатор частей тела. Второй слой принимает во внимание предполагаемые распределения классов первого и представляет собой перезагружаемый верхний участок для повторного соединения, моделируя взаимозависимость и взаимодействие деталей. Это приводит к созданию структуры оценки позы, которая учитывает зависимости между частями тела для уже локализации суставов и, таким образом, способна обойти типичные неоднозначности древовидных структур, таких как ноги и руки.

В экспериментах демонстрируется тот факт, что зависимые от частей тела регрессоры суставов достигают более высокой точности локализации суставов, чем современные методы, основанные на деревьях.

Для получения более точных результатов распознавания частей тела широко использовались сверточные нейронные сети (CNN), которые значительно повысили точность оценки. «Convolutional pose machines» [15], описанные в 2016 году также в работе специалистов Университета Карнеги-Меллона, обеспечивают систему последовательного прогнозирования для изучения неявных пространственных моделей. В этой работе показывается систематический расчет того, как сверточные сети могут быть включены в структуру PoseMachine для изучения особенностей изображения и

пространственно-зависимых моделей изображения для задачи оценки позы. Вклад этой статьи заключается в том, чтобы неявно моделировать долгосрочные зависимости между переменными в задачах структурированного прогнозирования, таких как оценка сочлененной позы. Это достигается путем проектирования последовательной архитектуры, состоящей из сверточных сетей, которые работают непосредственно на картах признаков из предыдущих уровней, производя все более и более точные оценки местоположений деталей, без необходимости явного вывода графического стиля модели. Данный подход направлен на устранение характерных трудностей исчезновения градиентов во время обучения, обеспечивая естественную целевую функцию обучения, которая обеспечивает промежуточный контроль, тем самым пополняя градиенты с обратным распространением.

Еще один вклад в исследование систем распознавания поз внесла группа исследователей Университета Бонна совместно с Google и Amazon в 2018 году. В статье «PoseTrack: a benchmark for human pose estimation and tracking» [6] исследователи обращают внимание на проблемы существующих систем для оценки позы на основе видео, которые пытаются добиться хороших результатов на реалистичных кадрах с несколькими людьми и часто не в состоянии вывести траектории тела, непротиворечивые во времени. Для решения этой проблемы в этой статье вводится PoseTrack, который представляет собой новый крупномасштабный эталон для оценки позы человека и сочлененного отслеживания на основе видео. Новый эталонный тест включает в себя три задачи, которые сосредоточены на 1) оценке позы одного человека в нескольких кадрах, 2) оценке позы нескольких людей в видео и 3) отслеживании сочленений нескольких людей. Чтобы установить эталонный тест, собирается, анализируется и выпускается новый набор данных, в котором представлены видео с участием нескольких человек, помеченных как персонализированные треки и четко сформулированные позы. Предоставляется общедоступный централизованный сервер оценки,

позволяющий исследовательскому сообществу оценивать результаты тестов. Кроме того, проводится обширное экспериментальное исследование последних подходов к отслеживанию сочлененных поз, анализируются сильные и слабые стороны уровня техники. В работе предполагается, что предлагаемый эталонный тест будет стимулировать продуктивные исследования, предоставляя большой и репрезентативный набор обучающих данных, а также предоставляя платформу для объективной оценки и сравнения предлагаемых методов.

На сегодняшний день существует ряд исследований, посвященных не просто распознаванию поз, а даже распознаванию суставов рук. Так в работе «Hand Keypoint Detection in Single Images using Multiview Bootstrapping» [13], написанной в 2017 году, описывается первая безмаркерная 3D-система захвата движений рук, способная реконструировать сложные взаимодействия рук и музыкальные исполнения без ручного вмешательства. Статья описывает подход, в котором используется многокамерная система для обучения мелкозернистых детекторов для ключевых точек, склонных к окклюзии, таких как суставы рук. В статье это называется начальной процедурной загрузкой: начальный детектор ключевых точек используется для создания шумных меток в нескольких видах руки. Обнаружение шумов затем триангулируется в трехмерное пространство с использованием многовидовой геометрии или помечается как выбросы. Наконец, спроектированные триангуляции используются в качестве новых помеченных обучающих данных для улучшения детектора. Этот процесс повторяется, генерируя больше помеченных данных в каждой итерации. В конечном счете достигаются результаты, аналитически связывающие минимальное количество просмотров для достижения целевых истинных и ложных положительных оценок для данного детектора. Этот метод используется для обучения ручного детектора ключевых точек для отдельных изображений. Полученный детектор ключевых точек работает в режиме реального времени на изображениях RGB и имеет точность, сопоставимую с методами, в которых используются датчики

глубины. Единый детектор вида, триангулированный по нескольким видам, позволяет захватывать движения 3D без маркера при сложных взаимодействиях объектов.

Рассмотренные исследования позволяют сделать вывод о том, что решение задачи распознавания людей на изображениях вызывает интерес исследователей во всем мире. Данная тема остается актуальной и по сей день. Однако, несмотря на множество методов распознавания, неразрешенным остается вопрос эффективного распознавания поз пешеходов в режиме реального времени в условиях окклюзий. Большинство вышеперечисленных работ показали хорошие результаты распознавания на снимках, где условия съемки приближены к идеальным: позы четко читаются, практически отсутствуют перспективные искажения. В данной работе мы использовали кадры с уличных камер видеонаблюдения, на которых мы можем видеть различного рода помехи: низкое разрешение картинки, перспективные искажения, перекрытия пешеходов другими пешеходами.

Глава 1. Данные и использованные инструменты

В первую очередь, следует рассказать об использованных в данной работе инструментах. В качестве языка программирования был выбран язык Python и среда разработки Jupyter Notebook. Жан Франсуа Пьюже специалист из отдела анализа данных IBM, основываясь на результатах поиска тенденций сайта indeed.com, выразил мнение, что сегодня Python является наиболее популярным языком для задач искусственного интеллекта и машинного обучения. Большой выбор библиотек - одна из главных причин, по которой Python является самым популярным языком программирования, в том числе и для задач анализа и обработки изображений. NumPy - это первый выбор среди разработчиков и исследователей данных. Это пакет Python, доступный для выполнения научных вычислений. Нами были использованы следующие библиотеки. С помощью NumPy вы можете использовать n-мерные объекты массива, функции для выполнения сложных математических операций, таких как преобразование Фурье, линейная алгебра, случайное число и т.д. TensorFlow - библиотека с открытым исходным кодом была разработана Google для вычисления графов данных с помощью мощных алгоритмов машинного обучения. Библиотека была разработана, чтобы удовлетворить высокий спрос на обучение нейронных сетей. Matplotlib - библиотека 2D-графиков для проектирования различных фигур в различных форматах, создания гистограмм, графиков, диаграмм рассеяния т.д. Существует несколько библиотек Python, которые обеспечивают надежную реализацию ряда алгоритмов машинного обучения. Одним из наиболее известных является Scikit-Learn, пакет, который предоставляет эффективные версии большого количества распространенных алгоритмов. Для создания клиент-серверного приложения был использован Flask - фреймворк для веб-разработки, это один из самых популярных модулей Python для веб-разработчиков.

Еще одним полезным инструментом стала библиотека алгоритмов компьютерного зрения обработки изображений и численных алгоритмов с

открытым кодом OpenCV. Библиотека разрабатывалась для таких языков как C, C++, Python, Java, Matlab и еще ряда языков программирования. Библиотека включает в себя алгоритмы идентификации объектов, возможность работы с мобильной робототехникой, распознавание лиц и жестов, отслеживание движения, сегментацию, SFM, AR, наивный байесовский классификатор и многие другие полезные пакеты, многие из которых оказались очень полезны в данной работе.

Источником данных послужили порталы (www.skylinewebcams.com, www.sochi.camera.ru, www.worldcam.eu) онлайн камер видео наблюдения со всего мира, откуда были собраны все видео материалы для анализа траекторий движения пешеходов. В частности, были взяты видео с камер наружного видео наблюдения городов Сочи, Нью-Йорк, Дублин, Миконос и камера с побережья Исландии. Всего было накоплено 7 видео по 10 часов, где были представлены пешеходные улицы при разных погодных условиях, на некоторых представлено как ночное, так и дневное время суток и помещения торгового центра.

Все данные обрабатывались на ноутбуке Acer с операционной системой Linux Ubuntu, процессором Intel Core i3-4005U 1.7ГГц, оперативной памятью 8Гб и графическим контроллером NVIDIA GeForce 840M 2Гб.

Следующим этапом после накопления данных стало планирование архитектуры системы. Это должно быть универсальное клиент-серверное приложение (Рис. 1.).

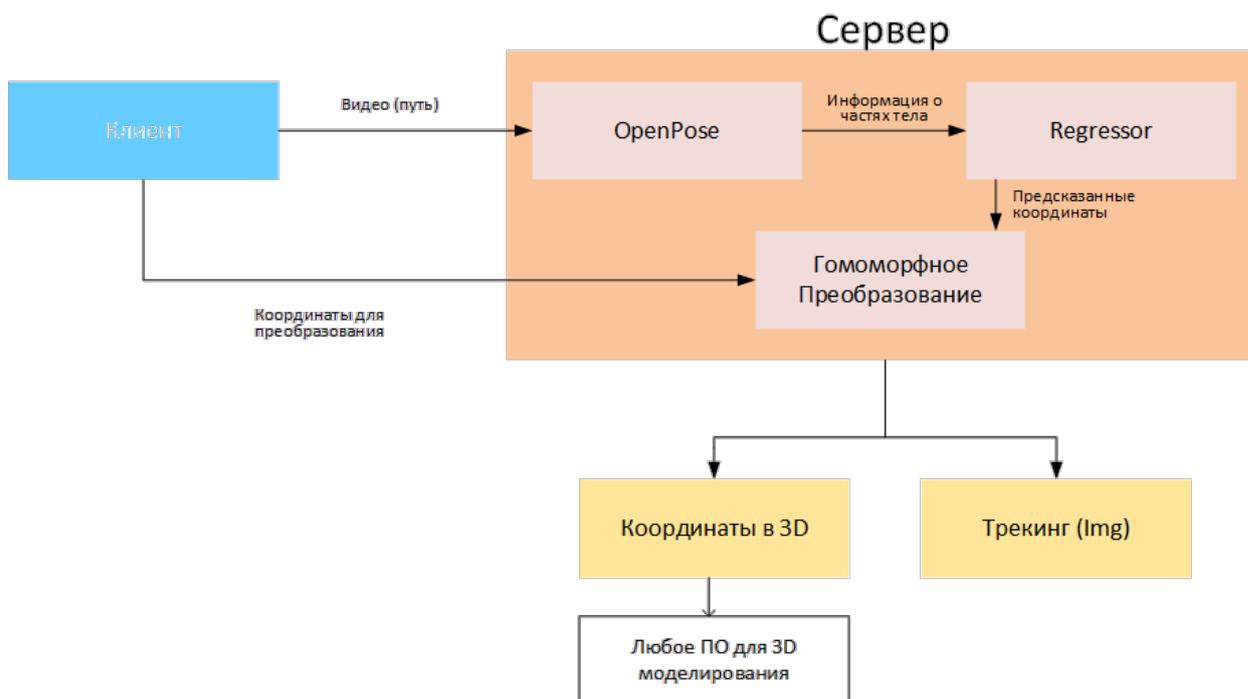


Рис. 1. Архитектура системы

Клиент посылает серверу post-запрос с указанием пути к анализируемому видео и дополнительной информацией для исследования и анализа трекинга. Сервер обрабатывает полученную информацию, используя ряд инструментов, и возвращает пользователю преобразованные координаты и изображение трекинга. Опишем подробнее инструменты, которые включает в себя сервер.

Глава 1.1 OpenPose

Система принимает в качестве входных данных цветное изображение размера $w \times h$ и создает двумерные положения анатомических ключевых точек для каждого человека на изображении. Работа алгоритма подробно описана в работе «OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields» [7]. Сеть с прямой связью предсказывает набор двумерных тепловых карт S местоположений частей тела и набор двумерных векторных полей L , которые кодируют расположение и ориентацию конечностей Part Affinity Fields (PAFs).

Тепловая карта - это матрица, которая содержит в себе информацию об уверенности сети в том, что определенный пиксель содержит определенную часть тела. 18 (+1) тепловых проиндексированных карт, связанных с каждой из частей тела. PAFs - являются матрицами, которые содержат информацию о положении и ориентации частей тела. Они возвращаются парами: для каждой части есть PAFs в направлении оси X и PAF в направлении оси Y . 38 PAFs, связанных с каждой из пар. Наконец, тепловые карты и PAFs анализируются с помощью жадного алгоритма, для получения двумерных ключевых точек для всех людей на изображении. Обнаружение происходит в три этапа. Сначала первые 10 слоев VGGNet [14] используются для создания карт функций для входного изображения. Далее используется многоступенчатая сверточная нейронная сеть с 2 ветвями, где первая ветвь предсказывает набор двумерных тепловых карт (S) местоположений частей тела (например, колено, колено и т. д.). Вторая ветвь предсказывает набор двумерных векторных полей (L). После каждого этапа прогнозы из двух ветвей, а также особенности изображения объединяются для следующего этапа.

После работы нейронной сети, возникает необходимость извлечь местоположения частей тела из тепловой карты с помощью локальных максимумов. После определения “кандидатов” на каждую из частей тела нужно соединить их, чтобы сформировать пары. Исследователи вводят

предположение, пусть для данного изображения найдены набор точек для шеи и набор подходящих точек для бедер. Для каждой точки шеи есть возможная ассоциация или связь точки с каждым из правых бедер. Есть полный двудольный граф, где вершины являются потенциальными точками, а ребра являются потенциальными соединениями. Нахождение наилучшего соответствия между вершинами двудольного графа является хорошо известной проблемой в теории графов, известной как задача присваивания. Чтобы решить ее, каждое ребро на графе должно иметь вес. С помощью линейного интеграла можно дать каждому соединению оценку, которая будет сохранена в взвешенном двудольном графе и позволит решить проблему назначения.

Взвешенный двудольный график показывает все возможные связи между потенциальными точками в обе стороны и содержит оценку для каждого соединения. Далее исследователи решают задачу поиска связей, которые максимизируют общий счет, то есть решают задачу назначения. Последний шаг - преобразование обнаруженных связей в финальные скелеты. В результате, получается набор человеческих скелетов, где каждый человек представляет собой набор частей, где каждая часть содержит свой индекс, свои относительные координаты и свою оценку точности обнаружения.

Авторы статьи разместили в общий доступ две модели - одна обучается на наборе данных для нескольких человек (МРП), а другая обучается на наборе данных СОСО. Модель СОСО дает 18 опорных точек, а модель МРП-15.

Глава 1.2. Преобразование координат

Для восстановления координат в трехмерное пространство и отслеживания трекинга пешеходов необходимо провести преобразование координат ног пешеходов. В обычной жизни мы часто сталкиваемся с перспективными искажениями. Когда смотрим на изображение и видим квадраты, которые не являются квадратами, или круги, которые не являются кругами. Данное явление происходит из-за того, что объект наблюдения расположен под некоторым углом к нашему глазу, видеокамере или другому аналогичному устройству. Именно поэтому, одним из важных является вопрос восстановления изображений, подверженных перспективному преобразованию. Мы можем определить проективное преобразование плоскости как любое отображение точек на плоскости, которое сохраняет прямые. Для решения задач преобразования трехмерного пространства и двумерной плоскости в матричном виде вводится понятие однородных координат.

В математике однородные координаты или проективные координаты, введенные Августом Фердинандом Мёбиусом, представляют собой систему координат, используемую в проективной геометрии, аналогично декартовым координатам в евклидовой геометрии. Реальная проективная плоскость может рассматриваться как евклидова плоскость с добавленными дополнительными точками, которые называются точками на бесконечности и считаются лежащими на новой линии, линии на бесконечности. Существует точка на бесконечности, соответствующая каждому направлению (численно определяемая наклоном линии), не формально определенная как предел точки, которая перемещается в этом направлении от начала координат. Параллельные линии в евклидовой плоскости пересекаются в бесконечной точке, соответствующей их общему направлению. Для точки (x, y) на евклидовой плоскости для любого ненулевого действительного числа Z тройка (xZ, yZ, Z) называется набором однородных координат для этой точки.

По этому определению умножение трех однородных координат на общий ненулевой коэффициент дает новый набор однородных координат для той же точки. В частности, $(x, y, 1)$ является такой системой однородных координат для точки (x, y) . Например, декартова точка $(1, 2)$ может быть представлена в однородных координатах как $(1, 2, 1)$ или $(2, 4, 2)$. Исходные декартовы координаты восстанавливаются путем деления первых двух координат на третью. Таким образом, в отличие от декартовых координат, одна точка может быть представлена бесконечным числом однородных координат.

Для перевода системы координат исходного видеоизображения в систему координат двумерной карты применялось преобразование гомографии. «Гомография — между точками, лежащими на двух прямых, а также между прямыми линиями, проходящими через две точки, можно установить такое однозначное соответствие, что каждой точке одной прямой будет соответствовать одна вполне определенная точка другой, а также в другом случае — прямой линии, принадлежащей к пучку прямых линий, проходящих через первую точку, будет соответствовать вполне определенная прямая пучка, проходящего через другую точку. Такое соответствие двух прямолинейных рядов точек, а также двух пучков называется гомографическим, или проективным» [2]. Другими словами, гомография – это обратимое преобразование плоскости, отображающее прямые линии в прямые линии. Имеющиеся кадры, снятые видеокамерой с некоторого ракурса, можно получить с помощью преобразования, примененного к искомому изображению, снятому с некоторой высоты видеокамерой, направленной вертикально вниз. Так как гомография - обратимое преобразование, то искомое изображение можно получить, вычислив примененную гомографию, обратив её и применив обратное преобразование к имеющемуся изображению (Рис. 2).

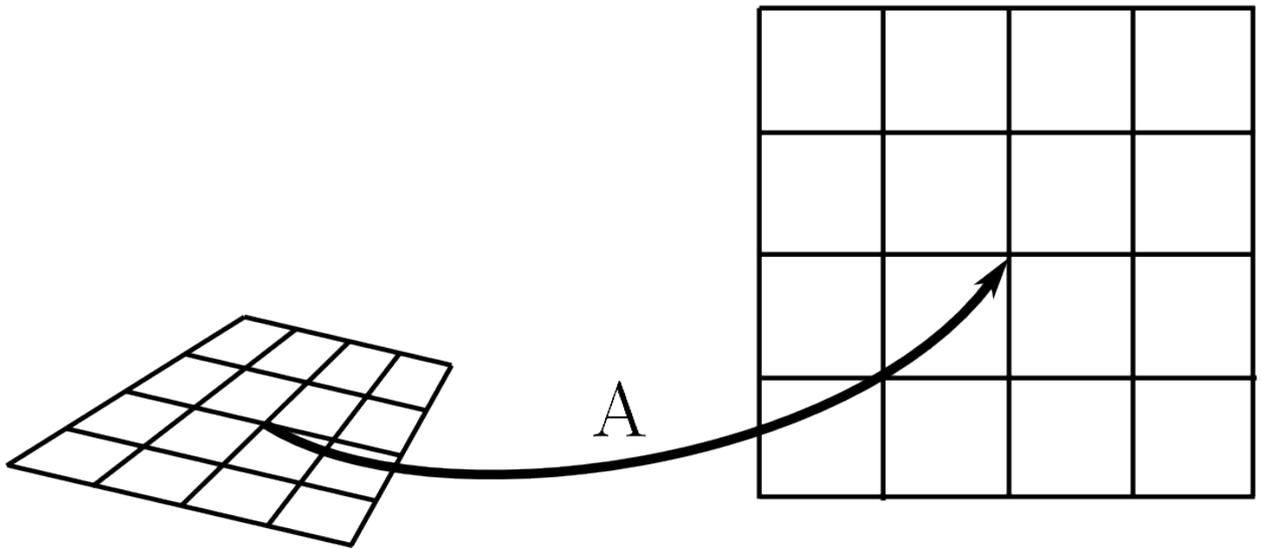


Рис. 2. Пример гомографии

Гомография представляется с помощью следующих параметров:

$$x' = \frac{(ax + by + c)}{(gx + hy + 1)}$$

$$y' = \frac{(dx + ey + f)}{(gx + hy + 1)}$$

где a, e — масштабирование по x и y соответственно; b, d - сдвиг по осям (вместе с a, e влияют на поворот); c, f - смещение по осям; g, h - изменение перспективы.

Матричный способ:

$$\bar{x}' = H\bar{x}$$

$$\text{где } H = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{pmatrix}$$

Итак, каждая точка нашего исходного кадра, и искомого имеет однородные координаты $(x, y, s)^T$, где x — координата по горизонтали, y —

координата по вертикали, s - ненулевой масштабный множитель, который и характеризует удаленность точки от наблюдателя. Преобразование плоскости находится путем решения системы линейных уравнений относительно координат исходного и искомого изображения.

Рассмотрим одно из таких уравнений:

$$a_i = H \times b_i$$

Поскольку мы используем однородные координаты, то вектора a_i и $H \times b_i$ совпадают по направлению, но не обязательно могут быть равны по длине из-за масштабного множителя. Тогда такое уравнение может быть записано через векторное произведение: $a_i \times H b_i = 0$.

Обратное преобразование задаётся матрицей, обратной к матрице гомографии: H^{-1} . Таким образом, точка a исходного кадра A соответствует точке $b = H^{-1} \times a$ искомого изображения B . Для преобразования требуется знать четыре опорные точки исходного изображения - углы какого-либо прямоугольного объекта, форма и расположение которого на изображении B хорошо известны.

Глава 1.3. Линейная регрессия

Для предсказания роста пешеходов в условиях окклюзий, было использовано два метода предсказания: линейная регрессия и многослойный персептрон.

В математической статистике линейная регрессия - это линейный подход к моделированию взаимосвязи между зависимой переменной и одной или несколькими независимыми переменными. Случай одной независимой переменной называется простой линейной регрессией. Для более чем одной зависимой переменной процесс называется множественной линейной регрессией:

$$Y = a + bx,$$

x - независимая переменная или предиктор. Y – зависимая переменная. Это значение, которое мы ожидаем для y (в среднем), если мы знаем величину x , т.е. это предсказанное значение y

- a – свободный член линии оценки или значение Y , когда $x=0$;
- b – угловой коэффициент или градиент оцениваемой прямой; он представляет собой величину, на которую Y увеличивается в среднем, если мы увеличиваем x на одну единицу; a и b - коэффициенты регрессии оцениваемо прямой.

Чтобы найти параметры, нам нужно минимизировать наименьшие квадраты или сумму квадратов ошибок. Конечно, линейная модель не идеальна, и она не будет точно предсказывать все данные, а это означает, что существует разница между фактическим значением и прогнозом. Ошибка легко вычисляется с помощью:

$$e_i = y_i - \hat{y}_i$$

Прогноз может быть выше или ниже истинного значения, что приводит к отрицательной или положительной разнице соответственно, поэтому мы

возводим ошибку в квадрат. Если возводим в квадрат, сумма ошибок может уменьшиться из-за отрицательных различий, но не из-за того, что модель хорошо предсказывает. Кроме того, возведение в квадрат ошибок учитывает большие различия, поэтому минимизация квадратов ошибок «гарантирует» лучшую модель.

Коэффициенты регрессии вычисляются следующим образом:

$$b = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

$$a = \bar{y} - \beta_1 \bar{x}$$

Глава 1.4. Многослойный персептрон

Многослойный персептрон (MLP) - это класс искусственной нейронной сети с прямой связью. MLP состоит как минимум из трех уровней узлов: входной слой, скрытый слой и выходной слой. За исключением входных узлов, каждый узел является нейроном, который использует нелинейную функцию активации. MLP для обучения использует контролируруемую технику обучения, называемую обратным распространением. Многослойность и нелинейная активация отличают MLP от линейного персептрона. Он может различать данные, которые не являются линейно разделимыми.

Если многослойный персептрон имеет линейную функцию активации во всех нейронах, то есть линейную функцию, которая отображает взвешенные входы на выход каждого нейрона, то линейная алгебра показывает, что любое количество слоев может быть уменьшено до двухслойной модели ввода-вывода. В MLP некоторые нейроны используют нелинейную функцию активации, которая была разработана для моделирования частоты потенциалов действия или запуска биологических нейронов. В обучении нейронных сетей самыми популярными функциями активации являются сигмоидальная функция, гиперболический тангенс и ReLU (Rectified Linear Unit).

Обучение происходит в персептроне путем изменения весов соединений после обработки каждого фрагмента данных на основе сравнения количества ошибок в выходных данных с ожидаемым результатом. Это пример контролируемого обучения, который осуществляется путем обратного распространения, обобщением алгоритма наименьших средних квадратов в линейном персептроне.

Ошибка представляется в выходном узле j в n -й точке данных как

$$e_j(n) = d_j(n) - y_j(n),$$

где d - целевое значение, а y - значение, создаваемое персептроном.
Вес узла корректируется на основе поправок, которые минимизируют ошибку во всем выводе, заданном

$$\varepsilon(n) = \frac{1}{2} * \sum_j e_j^2(n).$$

Используя градиентный спуск, изменение каждого веса

$$\Delta\omega_{j_i}(n) = -\eta \frac{\partial \varepsilon(n)}{\partial v_j(n)} y_i(n) \quad ,$$

где y_i - это выход предыдущего нейрона, а η - скорость обучения, которая выбрана для того, чтобы веса быстро сходились к ответу без колебаний.

Глава 2. Распознавание поз

Оценка двумерной позы человека - проблема локализации анатомических ключевых точек или «частей» - в основном сфокусирована на определении частей тела людей. Вывод поз множества людей на изображениях, особенно социально вовлеченных, представляет собой целый набор уникальных проблем. Во-первых, каждое изображение может содержать неизвестное количество людей, которые могут встречаться в любом положении или масштабе. Во-вторых, взаимодействия между людьми вызывают сложные пространственные помехи из-за контакта, окклюзии и суставов конечностей, что затрудняет ассоциацию частей. В-третьих, сложность во время выполнения, как правило, возрастает с увеличением количества людей в изображении, что затрудняет производительность в реальном времени. Поэтому, прежде чем научиться отслеживать траекторию движения всех пешеходов в режиме реального времени, необходимо обучиться детектировать самих пешеходов.

Применив данную систему, мы получили следующие результаты:



Рис. 3.1 Кадр улицы города Миконос (Греция).



Рис. 3.2 Обнаруженные позы.



Рис. 4.1. Кадр улицы города Дублин (Ирландия).

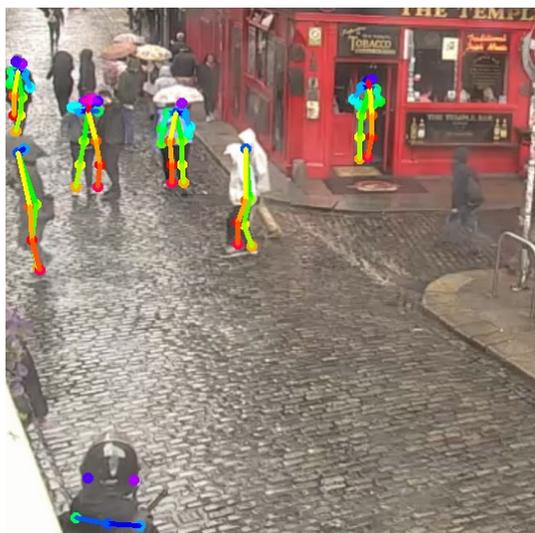


Рис. 4.2. Обнаруженные позы.



Рис. 5.1. Кадр побережья Исландии.



Рис. 5.2 Обнаруженные позы.

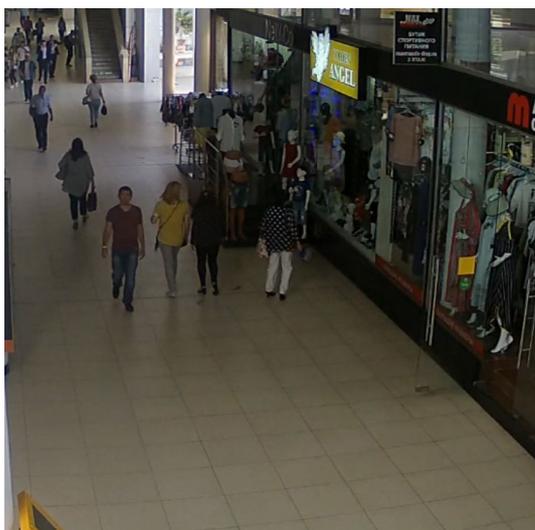


Рис. 6.1. Кадр с видео в торговом центре, май.



Рис. 6.2. Обнаруженные позы.

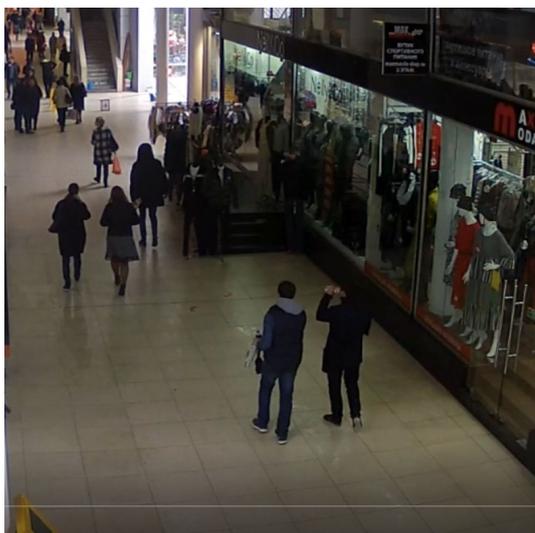


Рис. 7.1. Кадр с видео в торговом центре, декабрь.

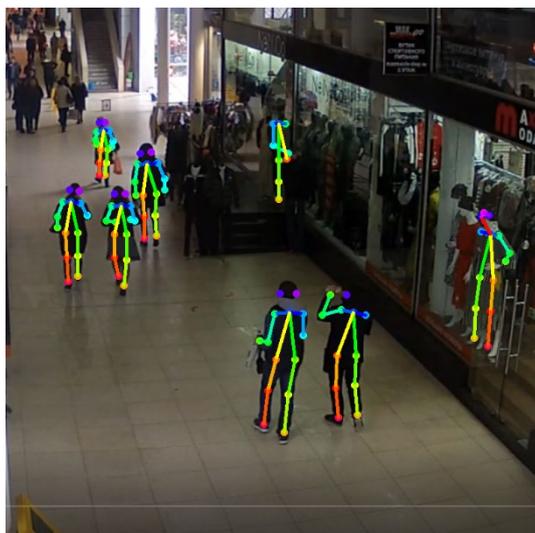


Рис. 7.2. Обнаруженные позы.

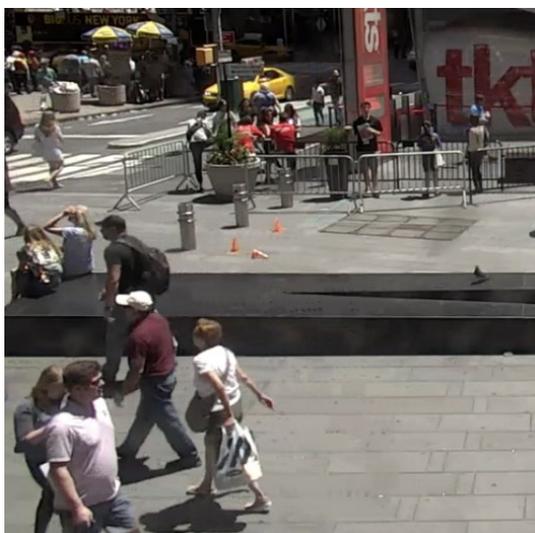


Рис. 8.1. Кадр улицы Таймс-сквер в городе Нью-Йорк (США)

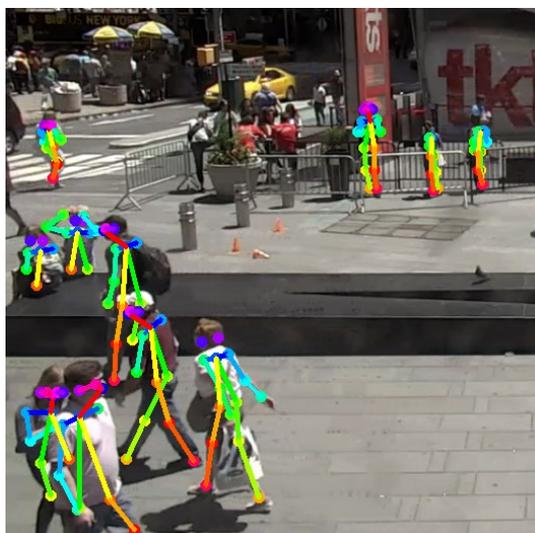


Рис. 8.2. Обнаруженные позы.

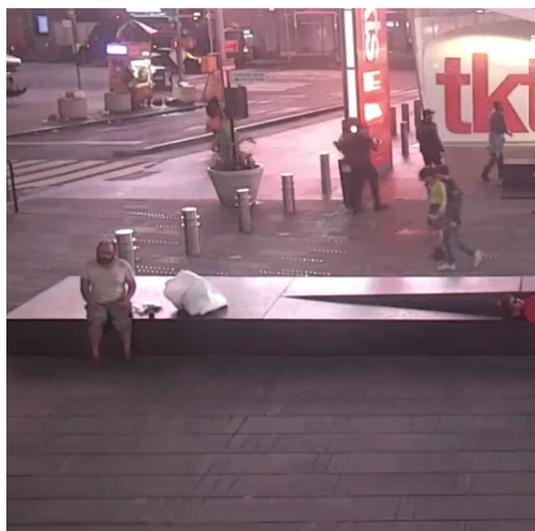


Рис. 9.1. Кадр улицы Таймс-сквер в городе Нью-Йорк (США). Ночь.



Рис. 9.2. Обнаруженные позы.

Исходный кадр имеет разрешение 368x368 пикселей. Для чистоты исследования параметры изображения (экспозиция, яркость, контрастность) остались неизменными. На распознавание одного кадра у модели OpenPose уходило до 0,30 секунд, что гораздо меньше скорости воспроизведения самого видео с частотой 30 кадров в секунду.

Как можно заметить по изображениям, система не всегда точно распознает людей в кадре. Всего можно рассмотреть два крайних случая: система не досчитала людей, и система обнаружила больше людей, чем есть в кадре. Под первую ситуацию попадают все рассмотренные случаи. Самой сложной ситуацией для распознавания стали ночные кадры на Рис 9.1 и кадры дождливой улицы на Рис. 4.1. Однако, в случае дождливой погоды, как на Рис4.1., несмотря на помехи в виде капель на объективе камеры и зонтов, системе удалось распознать большую часть пешеходов (Рис 4.2.) и даже тех, которые прятались под зонтами (см. Приложение, Рис 4.2., 6.2.). В режиме ночной съемке, несмотря на хорошую освещенность улицы Таймс Сквер, система не распознала пешеходов в тех местах, где она распознала их при дневной съемке (Рис 9.2.). Основной причиной, по которой модели OpenPose не удалось распознать некоторых людей является их большая удаленность от камеры. Система элементарно не может распознать в данном малочисленном скоплении пикселей человека. Этот результат можно объяснить тем фактом что, чем дальше человек расположен от камеры, тем меньше информации в виде пикселей мы можем получить. Даже человеческий глаз не всегда точно сможет сказать, что он видит на дальнем расстоянии. Еще одной причиной являются многочисленные окклюзии. Однако, нужно отметить, что в ряде случаев (см. Приложение, Рис 2.2.), несмотря на помеху в виде бетонного блока, система смогла распознать торговца у дороги. С другими результатами можно ознакомиться в Приложении.

Рассмотрим другую ситуацию, когда системе удастся обнаружить больше информации, чем есть на самом деле. На кадрах с торгового центра в витринах и у входов в магазины стоят манекены (Рис 7.1., Рис .6.1.). Система

периодически определяет их как людей, даже висящие на витрине плечики были распознаны как половина человека (см. Приложение Рис 10.2.). С другой стороны, это вряд ли можно считать это за ошибку в работе модели, учитывая тот факт, что манекены полностью повторяют строение человека.

В большинстве случаев система безошибочно определяет позы пешеходов и, по нашим статистическим подсчетам, определяет до 9 пешеходов из 12. Используя данную модель в качестве инструмента, важно помнить, что на вход принимается изображение относительно небольшого разрешения, что накладывает сильные ограничения на работу. Тем не менее, этот недостаток компенсируется точностью распознавания людей и частей их тела, даже в условиях низкого разрешения. Этот факт не трудно подтвердить наглядными примерами кадров, снятых в разных условиях.

Глава 3. Определение роста пешеходов в условиях потери данных

В ходе решения ранее поставленных задач, перед нами возникла проблема потери части данных. Как выяснилось, эта проблема была вполне ожидаема, так как на любых видео с камер видеонаблюдения, часть информации скрыта от наблюдателя из-за хаотичного движения пешеходов и перекрытия их друг другом. По-простому, мы теряли часть информации о положении ног, когда один человек в кадре шел впереди другого, тем самым закрывал его на половину или на треть. В промышленных масштабах работы приложения, прежде чем преобразовывать координаты или определять траекторию перемещения пешеходов системе, понадобится восстановить эти самые данные, используя даже малейшую информацию о наличии частей тела в пространстве. Поэтому, первые 5-6 минут система должна накопить имеющиеся данные, а затем обучиться дальнейшему предсказанию. За указанное время система накапливает информацию о 8-12 тысячах позах (в зависимости от многолюдности в кадре и времени суток на видео).

Одним из вариантов восстановления данных мог бы стать простой перебор вариантов: зная положение одной из ступней, коленей, бедер и т.д. дополнять данные о росте и, соответственно, о положении ступней путем прибавления определенного численного значения. Однако, перебор большого количества вариантов - далеко не самое оптимальное и гибкое решение, так как оно сильно зависит от целого ряда констант, которые мы не всегда можем иметь. Так как объем данных для обучения не является большим, нам не требовалось использовать систему для анализа больших данных. Было принято решение предсказывать недостающие данные о росте двумя методами: линейной регрессией и многослойным перспетроном.

Для реализации линейной регрессии нам понадобилось накопить полные данные о скелетах пешеходов, выдаваемые на выходе моделью OpenPose для каждого тестируемого видео: нормированные на единицу координаты опорных точек скелетов. Для предсказания использовались

координаты головы. Как показал небольшой эксперимент, модель OpenPose не всегда с вероятностью 100% выдает полную информацию обо всех опорных точках, но всегда на выходе выдает полную информацию об опорных точках головы пешеходов. Опираясь на полученное знание, мы использовали данные о положениях головы в качестве данных, относительно которых должен вычисляться рост людей на кадрах, чье положение ног скрыто от камеры впереди идущим человеком.

В данной работе мы также использовали многослойный персептрон со следующими параметрами: один скрытый слой, состоящий из 30 нейронов; функция активации: ReLU; количество итераций: 10000.

Обучение нейронной сети проводилось также с помощью накопленных данных о среднем расстоянии между головой и стопами пешеходов, чьи данные были полными. На выходе обучения мы получили координаты стоп пешеходов, потерянных по причине различного рода окклюзий. Важно отметить, что использование данной модели нейронной сети – временное и легко модифицируемое решение. В будущем мы можем переобучить сеть, используя информацию о других узлах, тем самым, возможно, получим более точное решение. Например, для предсказания роста можно использовать расстояние между опорными точками плеч, или длину руки, которые также несут в себе немало полезной информации. Учитывая соотношение между несколькими опорными точками скелетов одновременно, мы сможем эффективнее определять рост человека в кадре при разной удаленности пешехода от камеры видеонаблюдения.

Результаты работы линейного регрессора и многослойного персептрона приведены на графиках. Оси X и Y обозначают значение координат положения пешеходов на кадрах. Ось Z обозначает рост. Для наглядности на график выведена информация о росте 300 пешеходов. На первом графике (Рис.10.1) показан рост, исходя из полных данных полученных от модели OpenPose. На втором графике (Рис. 10.2.), показан рост, предсказанный с помощью линейной регрессии для пешеходов на кадрах, чьи данные о частях

тела представлены не полностью. Третий график (Рис. 10.3.) демонстрирует значение роста пешеходов, предсказанные многослойным персептроном.

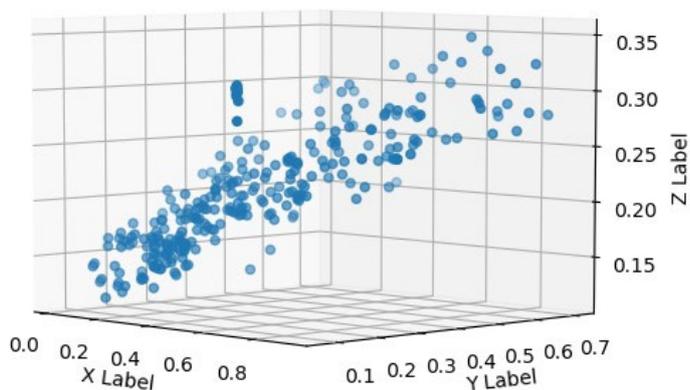


Рис. 10.1. График распределения роста пешеходов

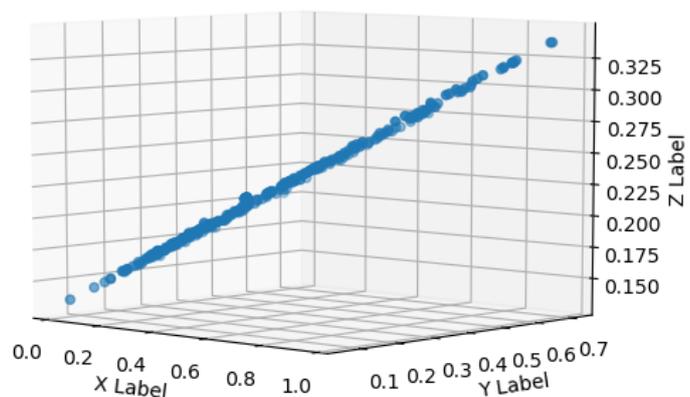


Рис. 10.2. График распределения роста пешеходов, предсказанного линейной регрессией

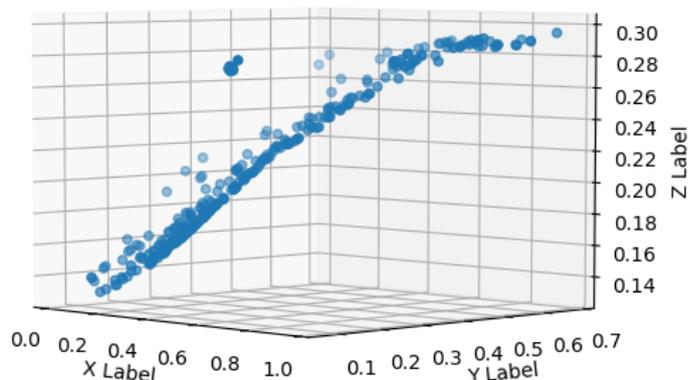


Рис. 10.3. График распределения роста пешеходов, предсказанного многослойным персептроном

Оценка работы методов проводилась с помощью метода среднеквадратичной ошибки MSE (Mean squared error). MSE оценивает

качество предиктора (т.е. функции, отображающей произвольные входные данные в выборку значений некоторой случайной величины), или оценочную функцию (то есть математической функции, отображающей выборку данных в оценку параметра совокупности из которого взяты данные). Определение MSE отличается в зависимости от того, описывает ли мы предиктор или оценочную функцию. Нас интересует качество работы предиктора, для это мы используем следующую формулу:

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2,$$

где Y_i – истинное значение, а \hat{Y}_i - предсказанное. Результаты работы методов представлены в Таблице 1.

Модель	Линейная регрессия	Многослойный перцептрон
MSE	0.00106757	0.00055808
Время выполнения, ms	8,28	10,4

Таб. 1. Сравнение работы линейного регрессора и многослойного перцептрона

Основным критерием работы методов является предсказание роста непосредственно на кадрах видео (Рис. 11.1, Рис. 11.2.).



Рис. 11.1. Рост предсказанный линейным регрессором



Рис. 11.2. Рост, предсказанный многослойным перцептроном

Как видно по кадрам, результаты предсказания практически не отличаются. Таким образом методы линейной регрессии и простая нейронная сеть стали быстрым, элегантным и дешевым способом предсказания роста пешеходов.

Глава 4. Перспективное преобразование координат

В данной работе мы используем кадры, снятые на одну статичную камеру, и потому не можем однозначно определить рост и положение частей тела пешеходов в трехмерном пространстве. Решение задачи восстановления роста требует дополнительных кадров, снятых с другого ракурса другой и, вероятно, не одной камерой. В рамках данной работы, нам требовалось восстановить непосредственно координаты ног пешеходов, которые являются единственной точкой пересечения тела и плоскости пола. Так как проективное преобразование включает в себя перемножение матриц, данный метод является простым, быстрым и дешевым способом трансформации координат.

В качестве опорных точек можно взять любые четыре точки, обрамляющие искомую преобразованную поверхность. Это может граница поверхности пола, тротуара, пешеходной зоны по которой перемещается основной поток пешеходов — любая область, которую пользователь хочет исследовать.

Полученная матрица гомографии была обращена и соответствующее преобразование применено к кадру видеозаписи с помощью функции `getPerspectiveTransform()` библиотеки `OpenCV`. Заметим, что гомографию достаточно задать лишь однажды. В дальнейшем можно пользоваться уже вычисленной матрицей преобразования, в случае, если ракурс видеокамеры, с которого снимается изображение, не меняется. В обратном случае, изображение, снятое видеокамерой, будет получено преобразованием, отличным от вычисленного, и применение обратного преобразования не даст желаемого результата. В данном методе, поиск точек для задания гомографии хорош и удобен тем, что может быть осуществлен автоматически; в этом случае при изменении ракурса съёмки система может пересчитать матрицу преобразований.

Используя метод гомографии, мы получили новые координаты пешеходов. Результаты преобразования можно показать наглядно на видео

или на отдельно взятых кадрах. Ниже представлены примеры преобразования кадров, с различных камер в разных условиях съемки. В правом столбце представлен исходный кадр, в левом – преобразованное поле.



Рис. 12.1. Кадр с видео в торговом центре, декабрь.



Рис. 12.2. Поле преобразованных координат



Рис. 13.1. Кадр с видео в торговом центре, май.

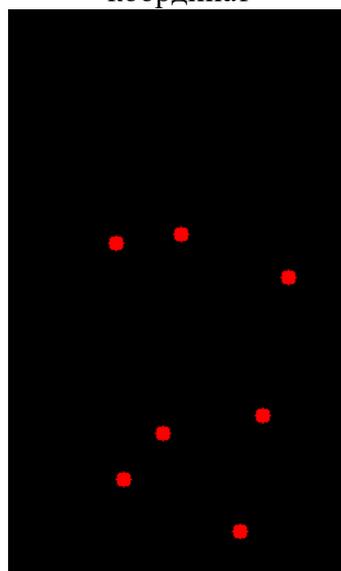


Рис. 13.2. Поле преобразованных координат

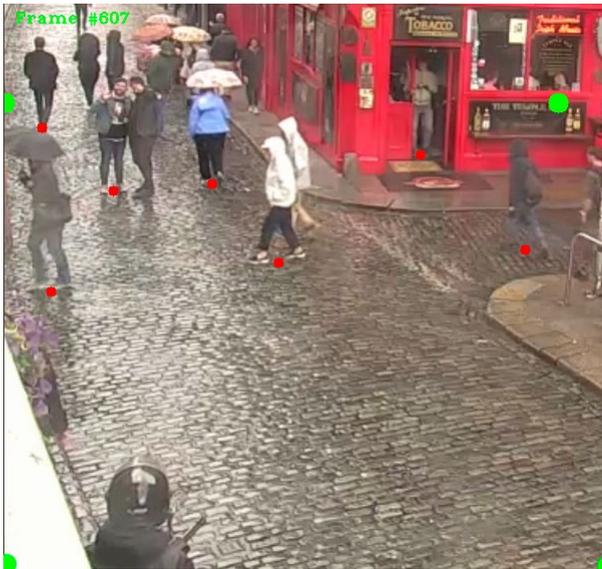


Рис. 14.1. Кадр с видео в торговом центре, декабрь.

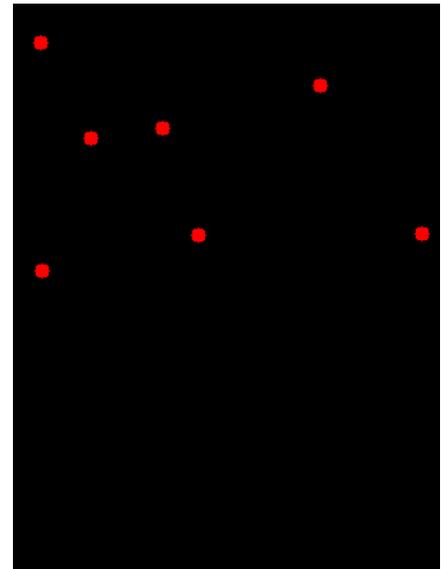


Рис. 14.2. Поле преобразованных координат



Рис. 15.1. Кадр улицы города Миконос (Греция).

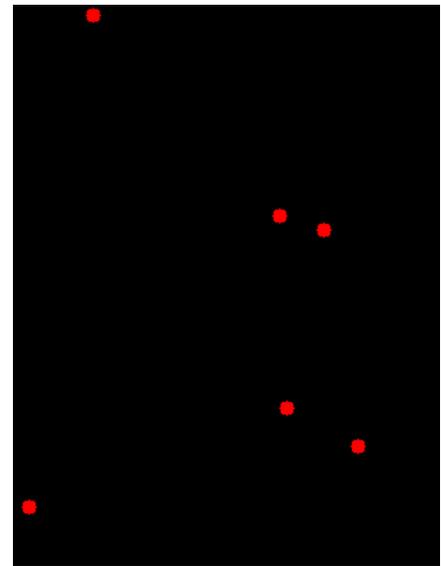


Рис. 15.2. Поле преобразованных координат



Рис. 16.1 Кадр улицы Таймс-сквер в городе Нью-Йорк (США)

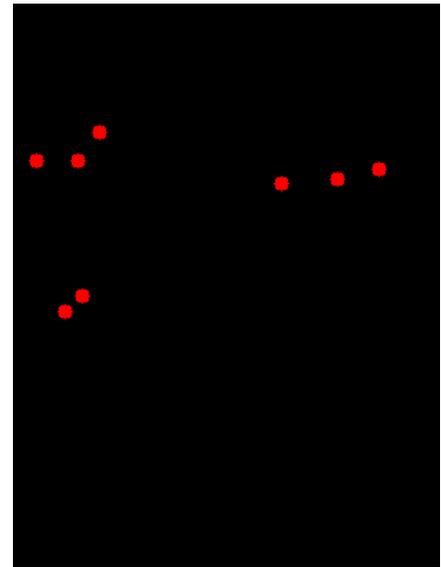


Рис. 16.2. Поле преобразованных координат



Рис. 17.1. Кадр побережья Исландии.

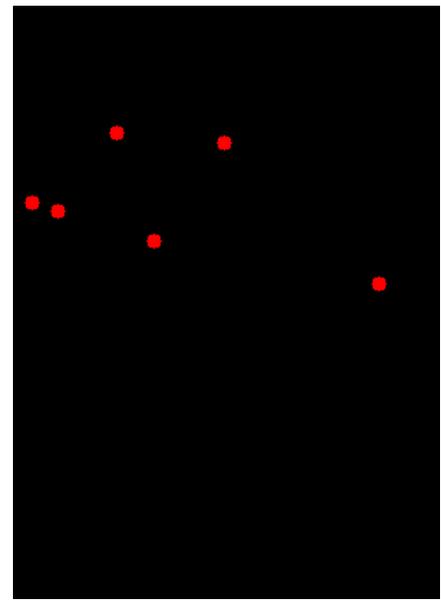


Рис.17.2. Поле преобразованных координат

Поле справа описывает “вид сверху” кадров с пешеходами. На обоих полях крупными красными точками отмечены положения ног. Зелеными точками обозначены опорные точки для преобразования, которые и образуют углы преобразованного поля. Новые полученные координаты в формате $(x,y,0)$ можно использовать для отображения положения пешеходов в трехмерном пространстве и отслеживания трекинга пешеходов. Данный метод никак не

зависит ни от количества пешеходов в кадре, ни от частоты или разрешения кадра. С помощью четырех точек, пользователь может сам задавать параметры области, которую хочет восстановить.

Вызов преобразования координат происходит с каждым обновлением кадра, и в среднем выполняется за 3,55 миллисекунды. Метод показал себя, как универсальный, точный и дешевый.

Получив, преобразованные координаты, мы смогли проследить трекинг пешеходов путем накапливания данных с каждого предыдущего кадра. В результате имеем решенную задачу оценки цифрового следа людей на улицах мира.

По результатам сбора информации о перемещениях пешеходов на улицах в разных странах мира, мы получили следующие изображения трекингов:



Рис. 18.1. Кадр улицы города Дублин (Ирландия).

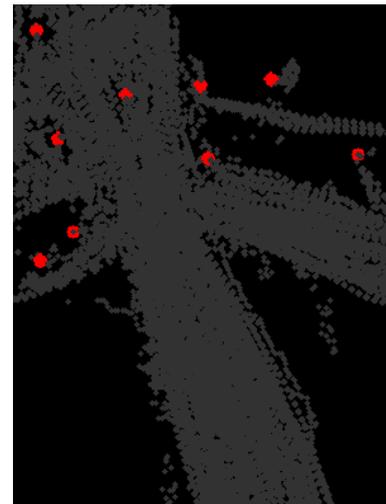


Рис.18.2 Трекинг пешеходов на улице Дублина



Рис. 19.1. Кадр побережья Исландии.

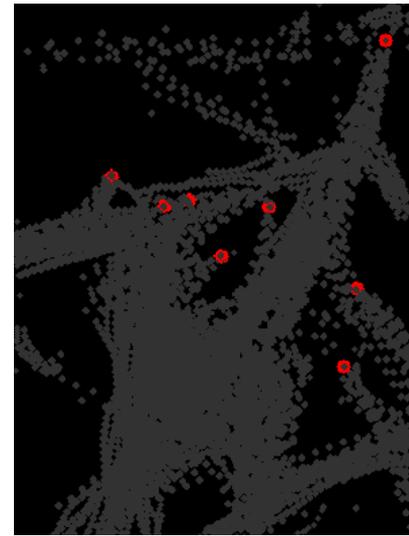


Рис. 19.2. Трекинг пешеходов на побережье Исландии

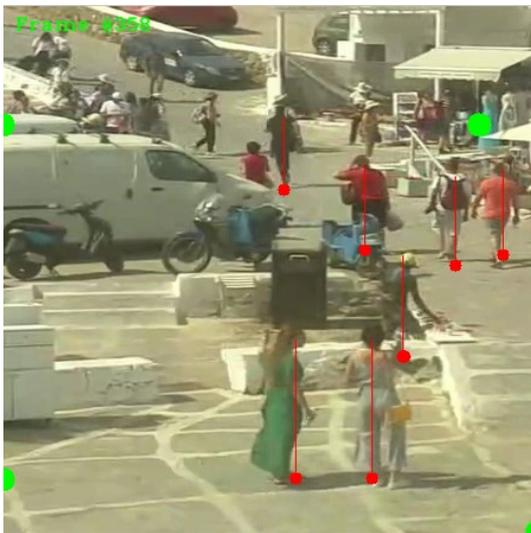


Рис. 20.1. Кадр улицы города Миконос (Греция).

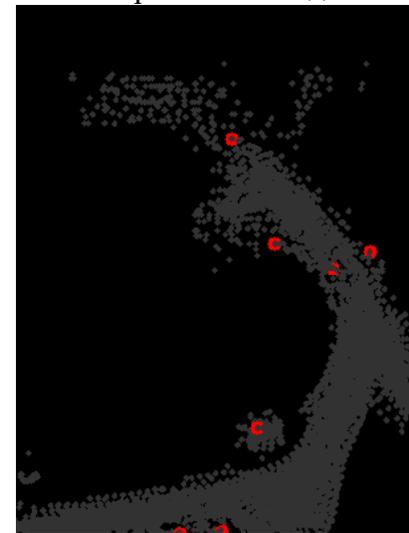


Рис. 20.2. Трекинг пешеходов на улице города Миконос



Рис. 21.1. Кадр с видео в торговом центре, декабрь.

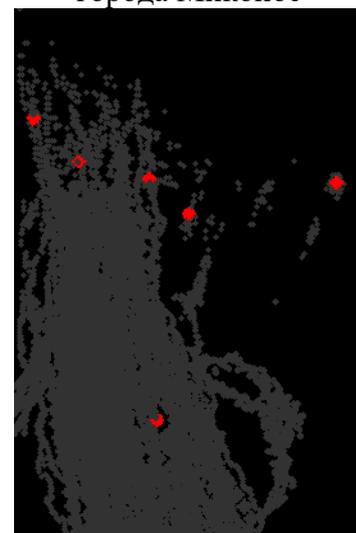


Рис. 21.2. Трекинг пешеходов в торговом центре



Рис. 22.1. Кадр с видео в торговом центре, май.

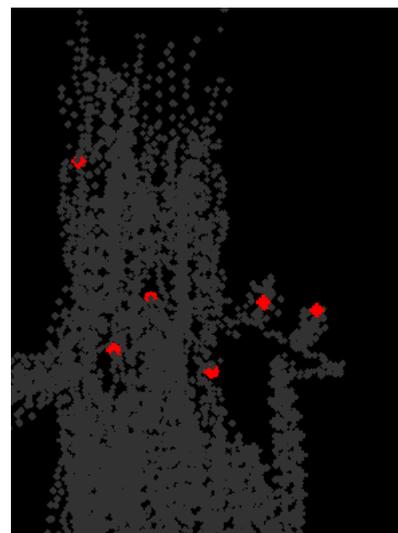


Рис. 22.2. Трекинг пешеходов в торговом центре



Рис. 23.1. Кадр улицы Таймс-сквер в городе Нью-Йорк (США)

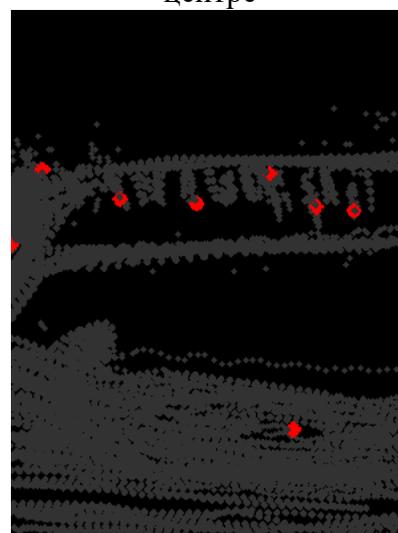


Рис. 23.2. Кадр с видео в торговом центре, май.

Серым обозначены пути перемещения, красные точки положения людей в кадре. Для получения трекинга понадобилось собрать данные о передвижениях пешеходов по улицах разных стран в разных условиях в течение 8-10 часов в зависимости от среднего количества человек в кадре. В результате мы получили четкую картину цифрового следа пешеходов, где четко прослеживаются траектории перемещения. В частности, на Рис 21.2. и Рис 22.2., отражающих информацию о поведении посетителей торгового центра, имеются ответвления от общего потокового трекинга. Обратив внимание на кадр, можно понять, что данное ответвление связано с наличием

входа в магазин. Похожую ситуацию можно наблюдать и на кадрах с торговой улицы Дублина (Рис 18.1.). Полученная информация является готовым маркетинговым инструментом, который мы можем уже сейчас использовать с целью моделирования полноценных виртуальных пространств на базе пересекающихся образов камер, для формирования модели поведения клиентов. Вопрос организации велосипедных дорожек также можно решить с помощью отслеживания трекинга пешеходов, что позволит оптимально разместить специальные пути для велосипедистов. Помимо этого, возможность отслеживать перемещения пешеходов может стать инструментом анализа кризисных ситуаций по поведению людей, например, в случае, когда затевается драка, или человеку на улице стало плохо.

Заключение

В данной работе мы ставили цель разработать систему анализа траекторий движения пешеходов методами искусственного интеллекта по данным камер видео наблюдения, способной работать в режиме реального времени. Для достижения этой цели было поставлено несколько задач, которые были успешно выполнены:

1. Были накоплены тестовых данных: видео, записанное статичной камерой видео наблюдения в разных погодных условиях, на улицах, в торговых центрах и в разное время суток;

2. Применена система, способной распознавать пешеходов на видео в режиме реального времени;

3. Определены координаты ног (x ; y) в двумерном пространстве;

4. Выполнено предсказание роста пешеходов в условиях потери данных;

4. Выполнено проективное преобразование координат для отслеживания трекинга пешеходов и перенесения данных из двумерного в трехмерное пространство;

5. Создано универсальное клиент-серверное приложение, возвращающего координаты пешеходов для отображения этих данных в трехмерном пространстве и изображение трекинга.

В результате, мы получили готовое клиент-серверное приложение, способное обрабатывать запрос пользователя (путь к видео и координаты исследуемой области) и возвращать полную информацию о трекинге пешеходов с кадров видео в границах запрашиваемой области и координаты для возможности восстановления трехмерной модели.

Список литературы

1. Ефимов, А.И. Алгоритм поэтапного уточнения проективного преобразования для совмещения изображений. Компьютерная оптика. Т. 40, №2. С. 258-265. 2016.
2. Граве Д. А. Гомография. Энциклопедический словарь Брокгауза и Ефрона: в 86 т. (82 т. и 4 доп.). - СПб., 1890 - 1907.
3. Малков А.Н., Михайлов И.А., Штерн Г.П., “Восстановление изображений, искажённых перспективным преобразованием” // Моделирование и анализ информационных систем, 16:2 (2009), 88–103.
4. Хартсхорн. Р. Основы проективной геометрии. — М.: Мир, 1970.
5. Andriluka M.; Roth S.; Schiele B. People-Tracking-by-Detection and People-Detection-by-Tracking // InProc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2008.
6. Andriluka M. PoseTrack: a benchmark for human pose estimation and tracking // InProc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2018.
7. Cao Zhe, Simon T., Wei Shih-En, Sheikh Y. Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields // InProc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) 2017.
8. Chen X., Yuille A. L. Articulated pose estimation by a graphical model with image dependent pairwise relations // In Advances in Neural Information Processing Systems, pages 1736–1744, 2014.
9. Chung Bryan WC. Pro Processing for Images and Computer Vision with OpenCV, 2017.
10. Dantone M., Gall J., Leistner C., Van Gool L. Human pose estimation using body parts dependent joint regressors // InProc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2013.
11. Hartley R., Zisserman A. Multiple View Geometry in Computer Vision, 2ed. Cambridge: Cambridge University Press, 2003.

12. Pudipeddi S., Chityala R. Image Processing and Acquisition using Python by, Publisher. Chapman and Hall/CRC, 2015.
13. Simon T., Joo H., Matthews I., and Sheikh Y. Hand keypoint detection in single images using multi view bootstrapping // InProc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2017.
14. Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition. University of Oxford, 2015.
15. Wei Shih-En ; Ramakrishna V. ; Kanade T. ; Sheikh Y. . Convolutional Pose Machines // InProc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2016.
16. Zhe Cao, Simon T., Wei Shih-En, Sheikh Y. OpenPose: Realtime Multi-Person 2D PoseEstimation using Part Affinity Fields // InProc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2011.

Приложение



Рис. 1.1. Кадр улицы города Миконос (Греция).

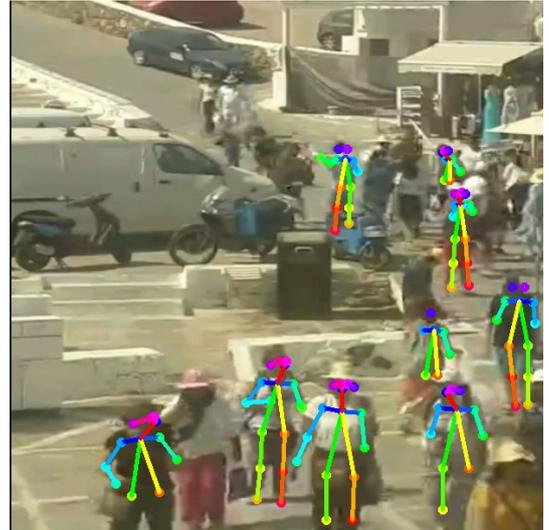


Рис. 1.2. Обнаруженные позы.



Рис. 2.1. Кадр улицы города Миконос (Греция).



Рис. 2.2. Обнаруженные позы.



Рис.3.1. Кадр улицы города Миконос (Греция).



Рис. 3.2. Обнаруженные позы.



Рис. 4.1. Кадр улицы города Дублин (Ирландия).



Рис. 4.2. Обнаруженные позы.



Рис. 5.1 Кадр улицы города Дублин (Ирландия).



Рис. 5.2. Обнаруженные позы.



Рис. 6.1. Кадр улицы города Дублин (Ирландия).



Рис. 6.2. Обнаруженные позы.



Рис. 7.1. Кадр побережья Исландии.



Рис. 7.2. Обнаруженные позы.



Рис. 8.1. Кадр побережья Исландии.

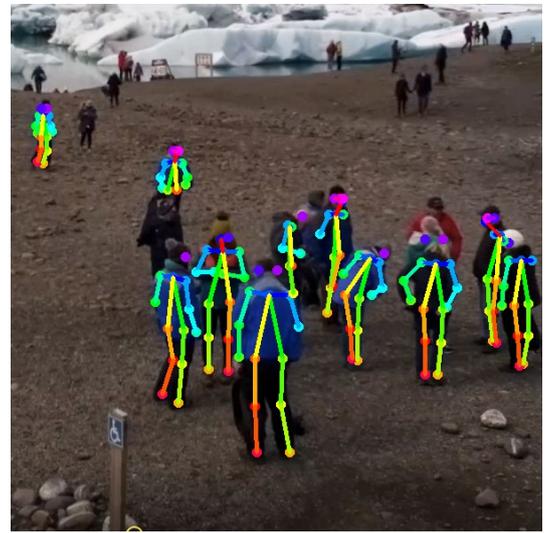


Рис. 8.2. Обнаруженные позы.

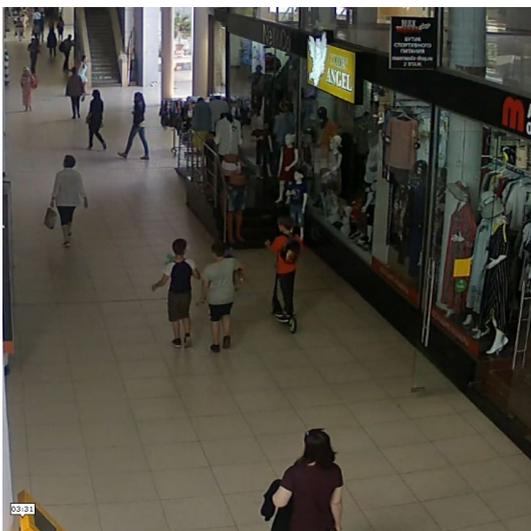


Рис. 9.1. Кадр с видео в торговом центре, май.



Рис. 9.2. Обнаруженные позы.



Рис. 10.1 Кадр с видео в торговом центре, май.



Рис. 10.2. Обнаруженные позы.

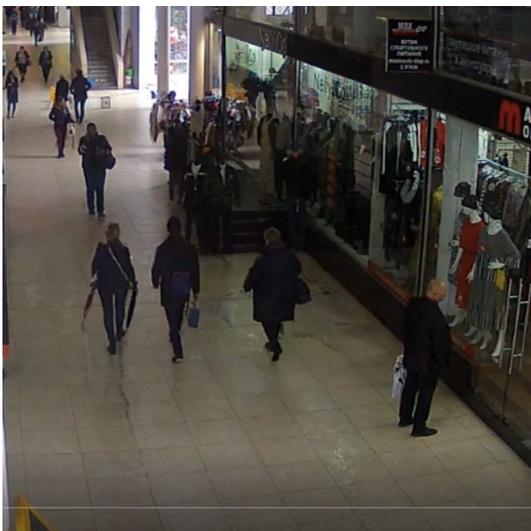


Рис. 11.1 Кадр с видео в торговом центре, декабрь.

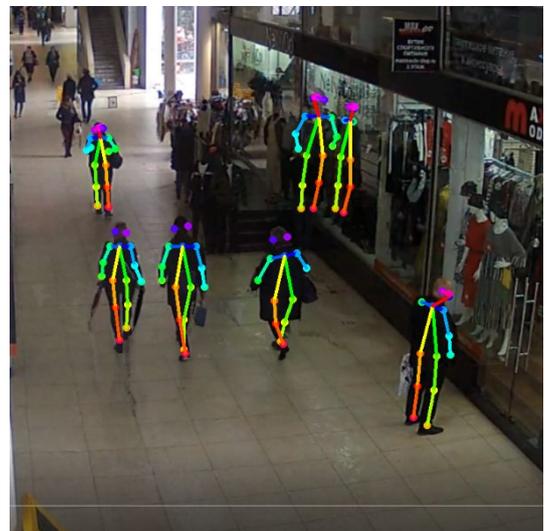


Рис. 11.2. Обнаруженные позы.

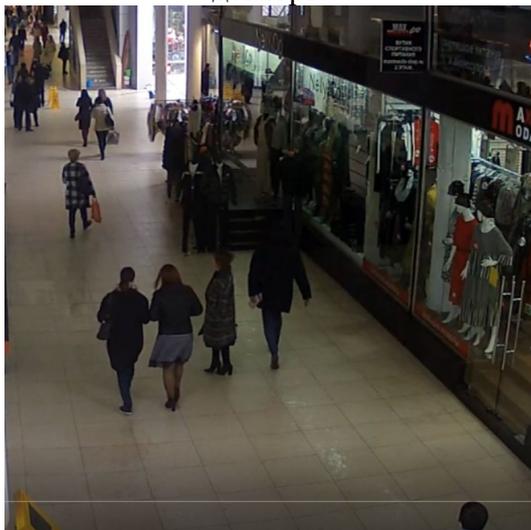


Рис. 12.1. Кадр с видео в торговом центре, декабрь.

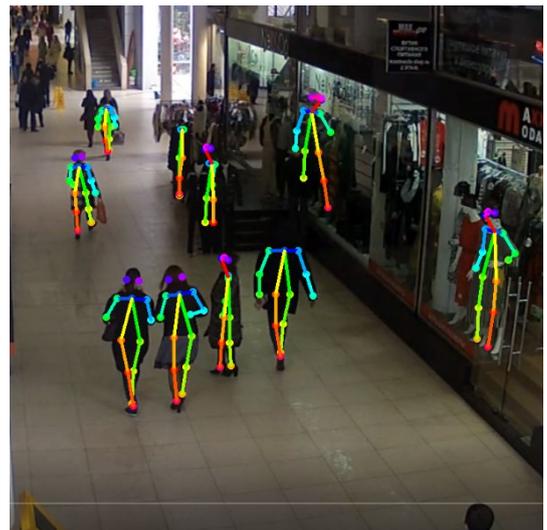


Рис. 12.2. Обнаруженные позы.



Рис. 13.1. Кадр улицы Таймс-сквер в городе Нью-Йорк (США)



Рис. 13.2. Обнаруженные позы.



Рис. 14.1. Кадр улицы Таймс-сквер в городе Нью-Йорк (США)



Рис. 14.2. Обнаруженные позы.

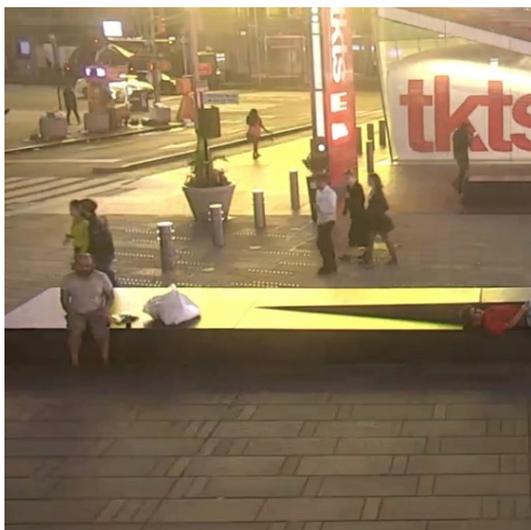


Рис. 15.1. Кадр улицы Таймс-сквер в городе Нью-Йорк (США). Ночь.



Рис. 15.1. Обнаруженные позы.

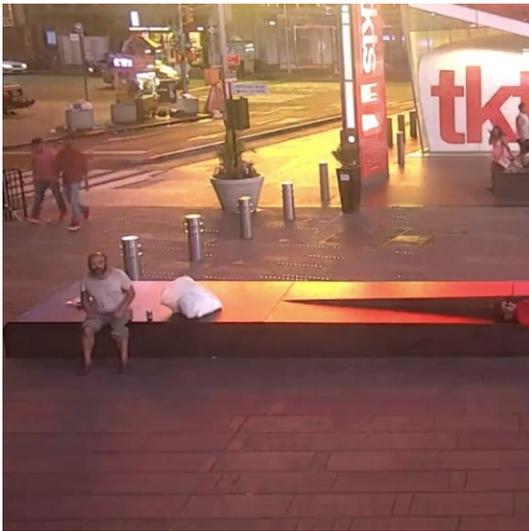


Рис. 16. 1. Кадр улицы Таймс-сквер в городе Нью-Йорк (США). Ночь.

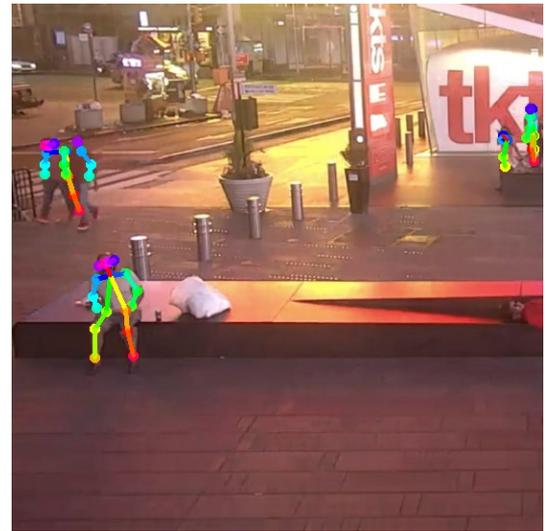


Рис. 16.2. Обнаруженные позы.