Основные принципы работы

Ситуационная видеоналитика обеспечивает детектирование определенных ситуаций в поведении объектов. Объектами, в данном случае, являются любые подвижные предметы в области обзора камеры видеонаблюдения. Это могут быть люди, автомобили, животные и другие предметы, движение и поведение которых необходимо идентифицировать. Прежде всего надо понимать, что анализируется последовательность кадров, поступающих от источника съемки – камеры. Такая последовательность называется видеопотоком.

Видеопотоки от камер видеонаблюдения

Современные камер виденаблюдения формируют последовательность кадров, кодируют ее и транслируют по сети потребителям. Существует множество форматов видео-данных, но в любом случае это последовательность кадров, которые передаются с определенной частотой. Кадром является изображение, характеризующееся разрешением и форматом (количеством бит на пиксель и их интерпретацией - какие биты за какую составляющую цвета отвечают). Внутри потока может быть использовано сжатие кадров для уменьшения объема передаваемых данных, но при показе на экране кадры всегда разжимаются до исходного состояния.

Алгоритмы видеоаналитики всегда работают с несжатыми кадрами. При этом, в один момент времени обрабатывается только один кадр, т.е. они обрабатываются последовательно. При обработке, важно знать, сколько прошло времени после предыдущего кадра. Эту величину можно вычислить из частоты кадров, но более практичным подходом является сопровождение кадра временной меткой -timestamp.

Таким образом видео поток характеризуется частотой кадров, их форматом и разрешением.

Общая схема работы анализа видеоизображения

Общая схема работы анализа видеоизображения представлена ниже на схеме.

Процесс делится на несколько последовательных этапов. На выходе каждого из них информация о происходящем в кадре дополняется все большими деталями. Также могут быть обратные связи между этапами, чтобы более тонко реагировать на изменения в кадре. Рассмотрим схему более детально.

Изменение размера и формата представления кадра

Первым этапом является подготовка кадра. Как правило, он значительно уменьшается в размерах. Так как в дальнейшей обработке будет участвовать каждый пиксель изображения. Соответственно, чем меньше кадр, тем быстрее все будет работать. Естественно, при сжатии теряется часть информации в кадре. Но это не критично, а даже полезно. Объекты, с которыми работает аналитика, в основном достаточно большие, чтобы не пропасть из кадра при сжатии. А вот всевозможный “шум”, связанный с качеством камеры, освещенностью, природными факторами, будет уменьшен.

Изменение разрешения происходит путем объединения нескольких пикселей исходного изображения в один. От способа объединения зависит та часть информации, которая будет сохранена.

Например, квадрат 3x3 пикселя исходного изображения должен быть преобразован в один пиксель результата. Можно просуммировать все 9 пикселей, можно взять сумму только 4-х угловых пикселей, а можно только один центральный. Приведем пример:

4 угловых пикселя:

Сумма всех пикселей:

Центральный пиксель:

Результат будет везде немного отличаться и по скорости и по качеству. Но иногда бывает так, что способ, теряющий больше информации, дает более ровную картинку, чем тот, который использует все пиксели.

Другим действием на данном этапе является изменение формата изображения. Цветные изображения, как правило, не используются, так как это тоже увеличивает время обработки кадра.

Цветное цифровое изображение определяется сочетание трех цветов: красного (Red), зеленого (Green) и синего (Blue). Каждый из них может изменяться в диапазоне от 0 до 255, т.е. один байт. Изображение формата RGB24 содержит 3 байта на каждый пиксель. В тоже время, черно-белое изображение (Y8) состоит из точек оттенков серого, и каждая из них содержит только один байт. При этом черно-белая картинка ненамного уступает в информативности цветной.

Перевод из цветного изображения в черно-белое осуществляется по формуле Y8=(R+G+B)/3. Для каждой точки сумма цветных байтов делится на три.

В результате получится то же изображение, но в градациях серого:

Фоновые модели

Самым важным этапом обработки является создание фоновой модели, т.е. определение изображения, на фоне которого двигаются объекты. Целью данного этапа является формирование фона сцены и получение разности между фоном и новым кадром. От алгоритмов этого этапа будет зависеть

качество работы всего алгоритма в целом. Если какой-то объект будет принят как фон или наоборот часть фона будет выделена как объект, то ошибки будут повторяться из кадра в кадр.

В дальнейшем будем рассматривать данный алгоритм на примере видеопотока, поступающего с камеры, которая установлена на железнодорожном вокзале.

В самом простом случае как фон можно взять кадр с пустой сценой:

Выберем кадр с движущимся объектом:

Если перевести оба кадра в Y8 (черно-белый) и вычесть из кадра с объектом кадр с фоном, то получим следующую разность:

Сделаем бинарное изображение, т.е. любая точка может быть 1 (белая) или 0 (черная). Заменим значение всех точек больших 0 на 255. В итоге мы перейдем к контрастному черно-белому изображению:

Вроде бы все хорошо, объект отделен от фона, имеет четкие границы. Но, во-первых, выделилась и тень объекта. А во-вторых, вверху кадра видны артефакты от шума изображения.

На практике такой подход никуда не годится. Любая тень, блик света, изменение яркости камеры испортят весь результат. **Именно в этом и заключается вся сложность задачи.** Объекты должны быть отделены от фона, при этом необходимо игнорировать природные факторы и шумы изображения: световые блики, тени от зданий и облаков, колеблющиеся ветки растений, артефакты сжатия кадров и т.д. При этом если вы ищите оставленный предмет, то наоборот, он не должен стать частью фона.

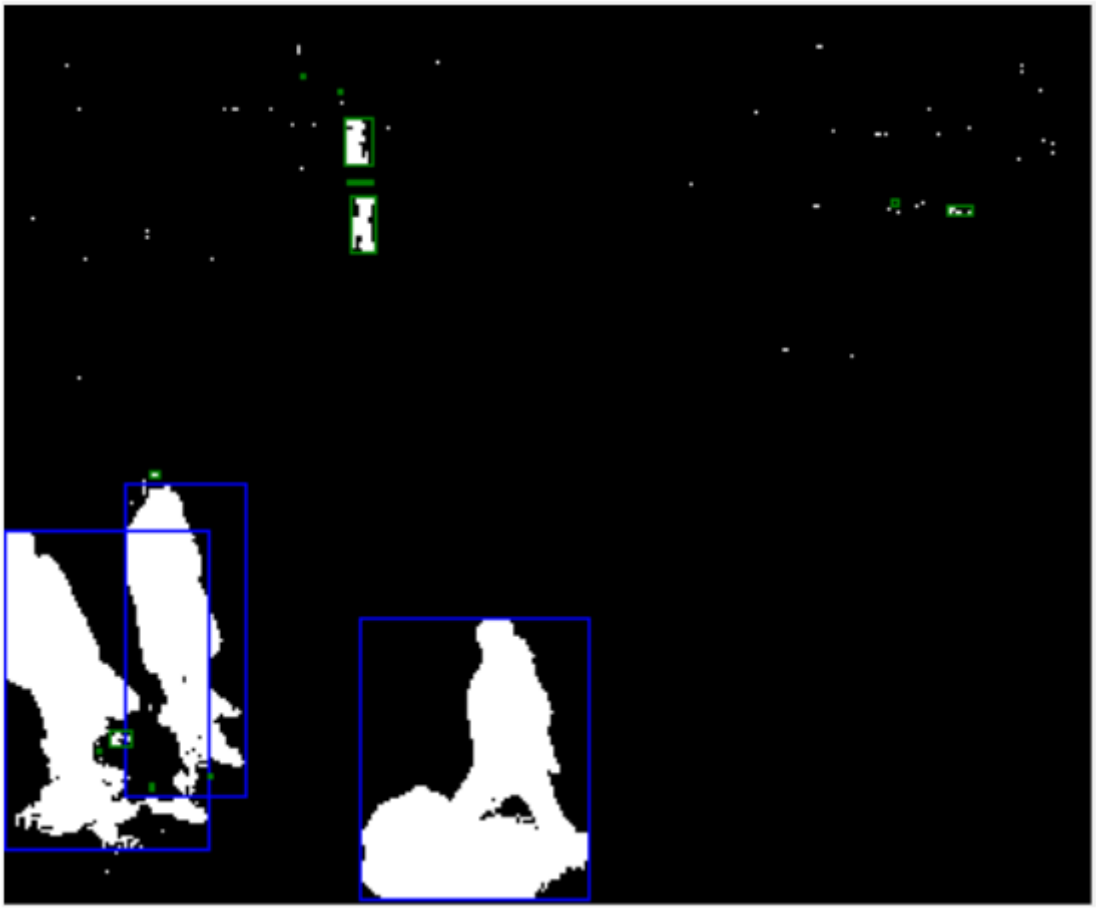
Существует множество алгоритмов, решающих эти проблемы с разной эффективностью. От простого усреднения фона до использования методов машинного обучения. Многие из них есть в проекте OpenCV (это активно развивающаяся кроссплатформенная библиотека компьютерного зрения, содержащая множество различных алгоритмов). Причем можно и комбинировать несколько подходов, что даст еще лучший результат. Но чем сложнее алгоритм, тем больше времени занимает обработка очередного кадра. При живом видео с чистотой 25 кадров в секунду у системы есть только 40 мс

на обработку. Поэтому выбор оптимального решения будет зависеть от задачи и ресурсов, выделенных на ее выполнение.

Формирование зон

Теперь нам надо выделить объекты, которые появились на фоновой модели. Именно для этого необходимо сделать разностный кадр, который будет отображать белые объекты на черном фоне:

Теперь нужно отделить объекты друг от друга и сформировать зоны, объединяющие пиксели объектов:



Тут сразу становятся видны все недостатки фоновой модели. Человек сверху разделен на несколько частей, много артефактов, тени от людей. Однако часть этих недостатков можно исправить на данном этапе. Зная область объекта, его высоту и ширину, плотность пикселей, можно отфильтровать лишние объекты.

На кадре сверху синими рамками указаны объекты, которые участвуют в дальнейшей обработке, а зелеными — отфильтрованные. Тут тоже возможны ошибки. Как видно, человек сверху, разделенный на несколько частей, тоже был отфильтрован из-за своих размеров. Эту проблему можно решить, например, за счет использования перспективы и предварительной юстировки камеры.

Возможны и другие ошибки. Например, несколько объектов могут быть объединены в один. Так что и на данном этапе есть большое поле для экспериментов.

Трекинг зон

На последнем этапе зоны превращаются в **подвижные** объекты. Тут используется результат обработки нескольких последних кадров. Главная задача – определить, что выделенная зона на двух соседних кадрах является одним и тем же объектом. Признаки могут быть самые разнообразные: размер, плотность пикселей, цветовые характеристики точек в зоне, предсказание направления движения и т.д. Именно здесь очень важны временные метки кадров. Они позволяют определить скорость объекта, предсказать где объект окажется в следующем кадре и рассчитать пройденное им расстояние.

На данном этапе можно исправить разовые ошибки предыдущего. Например, склеенные объекты можно разделить, учитывая их историю движения. С другой стороны, тут возможны проблемы. Самая главная из них - пересечение 2-х объектов. Частный случай этой проблемы, когда больший объект на длительное время заслоняет меньший.

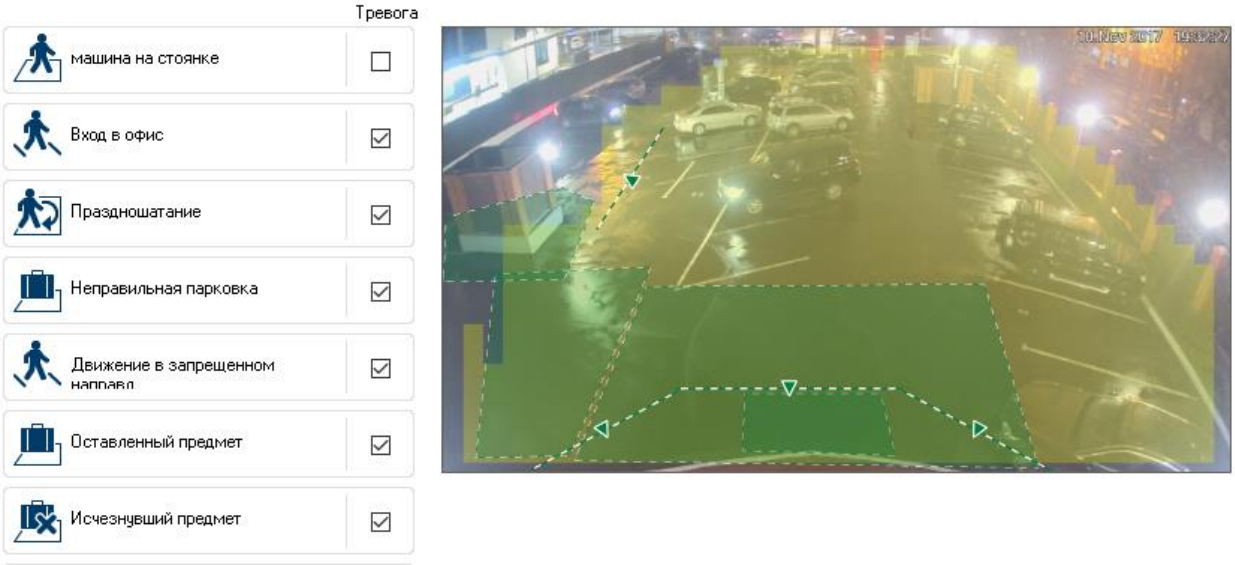
В алгоритме могут присутствовать обратные связи, улучшающие работу предыдущих этапов. Можно использовать информацию об объектах в сцене при формировании фона.

Например, можно отличить отставленный предмет и не делать его частью фона. Или бороться с “призраками”: если при создании фона на сцене был человек, то, когда он уйдет, на его месте появится объект-“призрак”. Понимая, что в этом месте начинается траектория движения объекта, можно быстро убрать “призрака” в фон.

Результатом работы всех этапов является список объектов в сцене. Каждый из них характеризуется размерами, плотностью, скоростью, траекторией, направлением движения и другими параметрами.

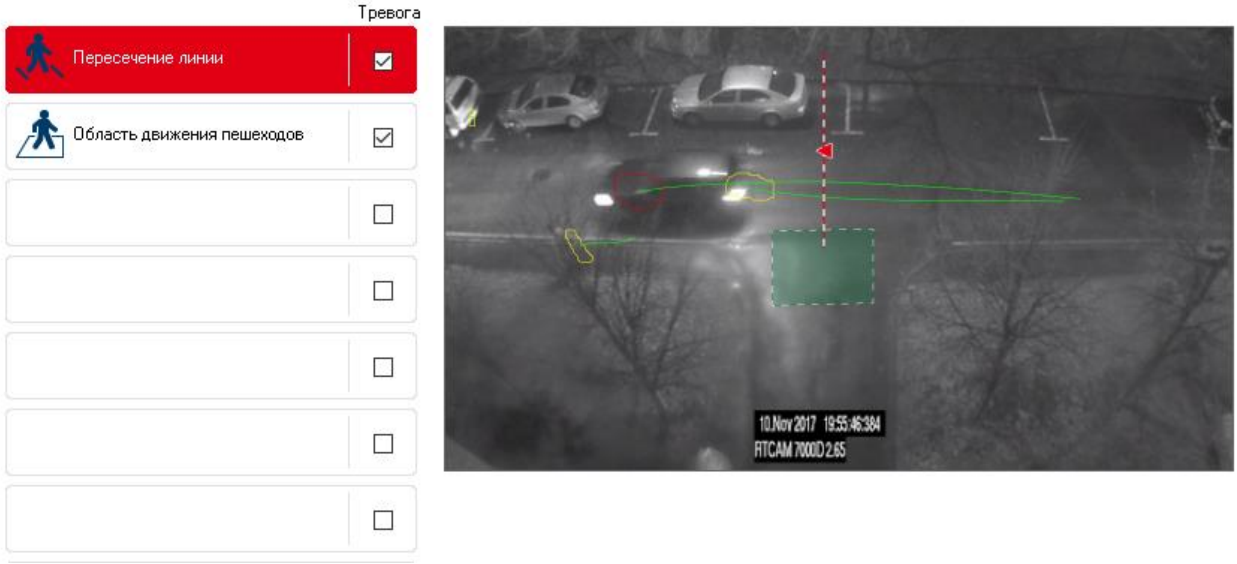
Этот список и используется при аналитике сцены. Можно определить пересечение объектом линии или движение в неправильном направлении. Можно подсчитать количество объектов в заданной зоне, праздное шатание, падение и множество других событий.

Приведем один из примеров использования видеоаналитики на камере Bosch. Ниже приведен интерфейс этой камеры, который демонстрирует настроенные на ней алгоритмы видеоанализа.



Из всего кадра выделена (желтым фоном) территория кадра, на которой осуществляется анализ подвижных объектов. В кадре определены зоны и рубежи, при пересечении объектами которых возникают тревоги.

На следующем примере такая же камера настроена на определение двух ситуаций: пересечение машины линии и вход пешехода в зону. В данном случае, мы видим, что определено событие связанное с проездом машины в указанном направлении, при этом, ели бы она двигалась в обратном направлении, то тревоги не было бы.



Области применения и их описание

Область применения ситуационной видеоаналитики на сегодняшний день очень широкая. Прежде всего она используется в системах обеспечения безопасности, а также при необходимости сбора статистических данных о перемещении и количестве тех или иных объектов. За каждым из применений стоит определенный метод обработки и анализа подвижных объектов в поле зрения камеры. При этом они анализируются по размеру, цвету траектории перемещения, а также по изменению фоновой модели.

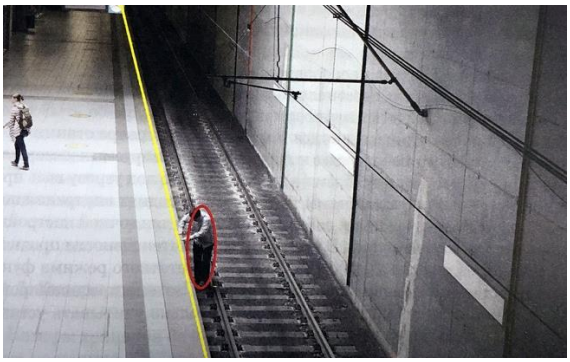
Рассмотрим эти методы подробнее.

Определение нарушений в перемещении объектов

Определение нарушений перемещения объектов является одним из самых распространенных методов видеоанализа. Он основан на анализе треков перемещения объектов, и в случае пересечения ими заранее определенных рубежей или входа в зоны, формирование тревожных сообщений.

Ниже пример показывает, как может осуществляться контроль линии безопасности на пироне в метро. Алгоритм работает так, если объект категории «человек» пересекает рубеж с лева на право, то возникает тревога.

Другой пример демонстрирует возможность контроля за перемещением по эвакуационным (пожарным) лестницам, которые часто используются при ограблении квартир. В этом случае алгоритм контролирует появление объекта категории «человек» в зоне кадра, которая очерчивает данную лестницу.



Определение оставленных предметов

Для транспортных узлов и других мест массово скопления людей важной задачей обеспечения безопасности является контроль за бесхозными (оставленными) вещами. Видеоаналитика здесь может помочь и автоматически проинформирует о появлении нового неподвижного объекта.

Конечно, в этом случае много ограничений. Результат зависит от мета установки камеры, ее разрешающей способности и интенсивности движения в кадре. В данном случае, метод основан на сравнивании фоновых моделей и появление в ней объекта, который до этого находился в движении.

Например: человек с чемоданов подошёл к стене вокзала и оставил его. Через некоторое время система фиксирует появление в фоне нового объекта и дает тревогу.

Частным случаем оставленного предмета является метод определения неправильной парковки автомобилей. При этом ожидается, что в фоновой модели появиться объект категории «автомобиль», т.е. движущийся до этого объект остановился и больше установленного времени неподвижен.

Подсчет объектов

На улицах и в торговых центрах нередко можно встретить людей, производящих разного рода подсчёты. Объектами подсчёта становятся, как правило, люди или автомобили. Примеры подсчётов могут быть: «Сколько было приобретено билетов в течение суток?» или «Сколько пассажиров проехало по 21-му маршруту в направлении центра после полудня?» или «Сколько людей сегодня зашло в помещение вокзала? Сколько из них осуществили посадку в поезда?» или «Сколько пассажиров зашло на паром? Сколько разместилось автомобилей на палубе?». На все эти вопросы может ответить видеоаналитика.

Самый точный из известных авторам способов подсчёта людей —«по головам», то есть при расположении камеры в зените. Высота подвеса камеры при этом обычно составляет не менее трёх метров, а направлена камера должна быть вертикально вниз. Алгоритм нуждается лишь в нескольких последовательных кадрах, чтобы обнаружить объект класса «человек». Счётчик срабатывает при пересечении виртуальной линии.

Для планирования развития автомобильной дорожной сети необходимо собирать статистику на въездах и съездах с трас. Также необходимо измерять интенсивность движения транспорта на городских улицах для оперативного изменения работы светофоров и других средств управления движением. Видеоаналитика может это обеспечивать при условии правильного размещения камер. Более точный результат достигается установкой камеры над каждой полосой движения и подсчет пересечения объектов (машин) такой же виртуальной линии, как в случае подсчета людей