

Лабораторная 4. СЕГМЕНТАЦИЯ ДВИЖУЩИХСЯ ОБЪЕКТОВ НА ОСНОВЕ ОЦЕНКИ ЭНЕРГИИ ДВИЖЕНИЯ

При разработке практически любой системы компьютерного зрения возникает задача автоматической сегментации целевых объектов. Под сегментацией в обработке изображений понимают разделение изображения на области, для которых выполняется определенный критерий однородности. Например, однородность может пониматься, как близость элементов разложения изображения по яркости или цветовому тону. Более конкретизированный термин «сегментация объекта» подразумевает разделение объекта и фона, иначе говоря – выделение объекта на изображении, или в видеоряде.

Если исходными данными являются изображения, то сегментацию объектов проводят по признакам яркости, цветности, локальной детальности (сгусткам высокочастотной энергии) и формы. Для оценки формы применяют аппарат контурного анализа, а оценку детальности традиционно осуществляют с помощью меры Розенфельда-Троя.

Дополнительно, для решения задачи сегментации объектов используют текстурные и цветовые признаки, а также более сложные техники, основанные на применении метода локальных бинарных шаблонов, Виолы-Джонса, ориентированных градиентов и других алгоритмов из области машинного обучения.

В видеоаналитике часто встречается задача, в которой исходные данные представляют собой видеопоследовательность, а целевые объекты перемещаются в поле зрения камер системы, то есть обладают признаком движения. Учет этого признака обеспечивает устойчивую сегментацию областей изображения, принадлежащих движущимся объектам на неподвижном, в том числе сложном, фоне.

При разработке алгоритмов сегментации объектов с учетом признака движения необходимо учитывать характер движения. Область интереса может обладать жестким (ригидным) и нежестким движением. Жесткое предполагает перемещение всех элементов объекта, в соответствии с основным направлением движения. Примерами являются движущийся автомобиль,

летательный аппарат, корабль. Нежестким движением обладают, например, идущий человек, летящая птица.

Достаточно простой и эффективный метод сегментации объектов на основе признака движения основан на оценке «энергии движения». Под этим термином понимают межкадровую разность телевизионных сигналов. В случае двух цифровых изображений, представляющих соседние (или отстоящие друг от друга на некоторое время t кадры, оценка движения может быть найдена как поэлементная абсолютная разность двух кадров:

$$MAD(x, y) = |L(x_j, y_i, t) - L(x_j, y_i, t - 1)|,$$

где, $L(x, y, t)$ – значение яркости пикселя с координатами (x, y) в кадре t , $L(x, y, t-1)$ – значение яркости пикселя с координатами (x, y) в кадре $t-1$.

Препарат межкадровой разности двух кадров представлен на рисунке 4.1.



Рис. 4.1. Пара кадров (верхний ряд) и их разность (внизу).

Очевидно, что в областях, где присутствует движение, абсолютная разность яркостей пикселей будет отлична от нуля, а там, где движение отсутствует будет близка к нулю.

Следует учитывать, что с помощью видеокамеры регистрируется только видимое двумерное движение объекта, которое является проекцией реального движения объекта в трехмерном пространстве на плоскость изображения. Видимое движение порождено изменениями яркости пикселей $L(x_j, y_i, t)$ с координатами x и y в кадре t по отношению к кадру $t-1$. Однако, ненулевой результат межкадровой разности не эквивалентен только видимому движению, так как, например, в статической сцене с изменяющейся освещенностью, где на самом деле движения нет, или может быть вызван шумами. К ненулевой межкадровой разности приведет и наличие глобального движения в сцене, которое возникает в случае движущейся камеры. При анализе видеоряда и сегментации объектов глобальное движение следует предварительно компенсировать.

При отсутствии глобального движения алгоритм сегментации объектов на основе энергии движения состоит из следующих шагов:

1. получение препарата межкадровой разности;
2. формирование контурного препарата для текущего кадра t ;
3. бинарное квантование контурного препарата и межкадровой разности;
4. объединение двух полученных препаратов с помощью операции логического «И» – получение изображения, содержащего оценку энергии движения, относящейся только к текущему кадру;
5. морфологическая фильтрация полученного изображения;
6. формирование строба объектов.

Второй и третий шаги необходимы потому, что изображение межкадровой разности отражает энергию движения, как предыдущего, так и текущего кадров. В результате центр области движения не совпадает с центром движения реального объекта. Определение контуров объектов в текущем кадре t и объединение препаратов по «И» устраняет этот фактор.

Для выделения контуров в текущем кадре t видеопоследовательности могут применяться различные алгоритмы (оператор Собеля, разницы гауссиан

и так далее). Традиционным в данном случае является использование оператора Собеля. Алгоритм использует ядра (маски) размером 3 на 3 элемента для выделения горизонтальных и вертикальных контуров соответственно:

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}.$$

Для изображения (кадр t) осуществляют свертку с каждым из ядер для вычисления приближённых значений производных по горизонтали G_x и по вертикали G_y . Приближённое значение величины градиента G можно поэлементно вычислить путём использования полученных приближенных значений производных по направлениям:

$$G(x, y) = \sqrt{G(x, y)_x^2 + G(x, y)_y^2}.$$

Пример контурного препарата, полученный для кадра t на рисунке 1 приведен на рисунке 4.2.

Контурное изображение $G(x, y)$ затем бинарно квантуют так же, как и изображение межкадровой разности $MAD(x, y)$, что дает возможность получить объединенное изображение $K(x, y)$, выполнив поэлементно логическую операцию «И»:

$$K(x, y) = MAD(x, y) \& G(x, y).$$

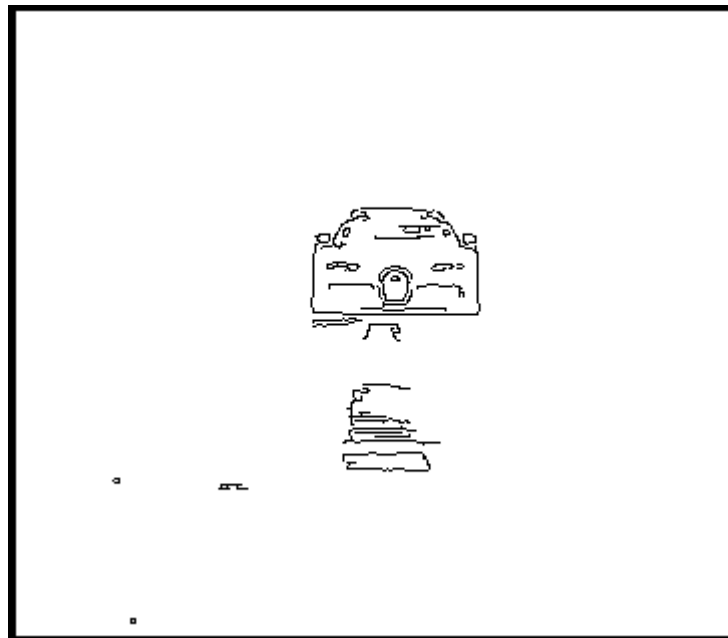


Рис. 4.2. Контурный препарат

В результате получают изображение (рис. 4.3), в котором выделенные движущиеся объекты занимают положение, соответствующее текущему кадру.

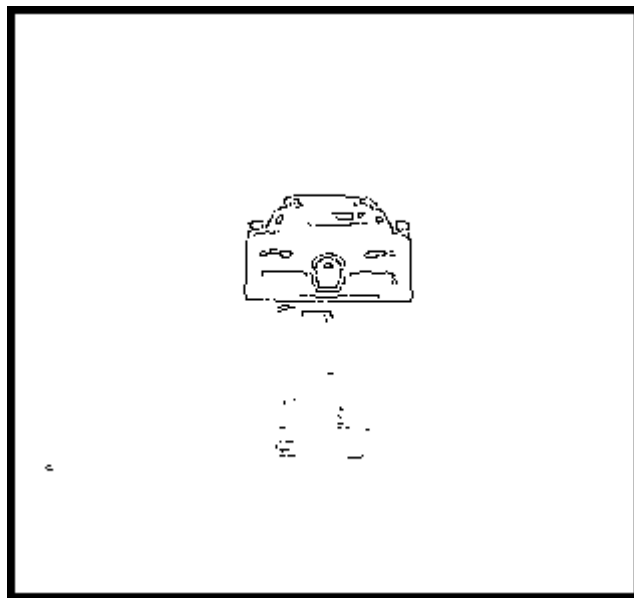


Рис. 4.3. Выделенный слой движения.

Полученный таким образом препарат $K(x,y)$ может содержать множество нежелательных фрагментов, вызванных, например, шумами. Для снижения помех целесообразно провести предварительную фильтрацию исходного изображения (например, медианным фильтром), а затем изображения межкадровой разности.

Изображения движущихся объектов на результирующем изображении $K(x,y)$ могут иметь разрывы, лакуны и прочие дефекты, которые затруднят точную сегментацию. Для удаления этих дефектов используют так называемую морфологическую фильтрацию, основанную на логических операциях. Морфологическую фильтрацию применяют к цифровым изображениям как бинарно квантованным, так и полутоновым.

Логические операции совершают между маской S и областью цифрового изображения, выделенной этой маской. Результат операции устанавливают в новую битовую матрицу на место, где находится фокус маски. Используют маски различной формы и размерности.

Морфологическая операция «открытие» (последовательное применение операций эрозии и наращивания) с маской малого размера (относительно предполагаемого размера объекта) позволяет подавить импульсный шум на изображении. Применение после этой операции «закрытие» (последовательное применение операций наращивания и эрозии) с увеличенным размером маски способствует удалению разрывов и заполнению лакун в изображениях объектов.

На заключительном шаге алгоритма организуют наложение строка на объект (рис. 4.4). Для формирования строка в методе сегментации объектов на основе «энергии движения» используют алгоритм проекций. Для результирующего изображения $K(x,y)$, прошедшего морфологическую фильтрацию, в каждом столбце и в каждой строке подсчитывают количество ненулевых («белых» в рассматриваемом примере) пикселей. По полученным данным строят соответствующие гистограммы, сопоставляющие номеру строки (столбца) – число ненулевых пикселей.

Затем выделяют проекции – совокупность строк (столбцов), у которых число ненулевых пикселей превышает некий порог и которые образуют на гистограмме непрерывный отрезок. Проекции, длина которых меньше заданной величины (определяемой априорным знанием о предполагаемых размерах объектов), отбрасывают, чтобы исключить изолированные помехи. На базе оставшихся проекций формируют строки.

Если в кадре движутся несколько объектов, то для установления соответствия между проекциями по горизонтали и вертикали используют правило: число пикселей в горизонтальной и вертикальной проекции одного объекта должны быть равны.

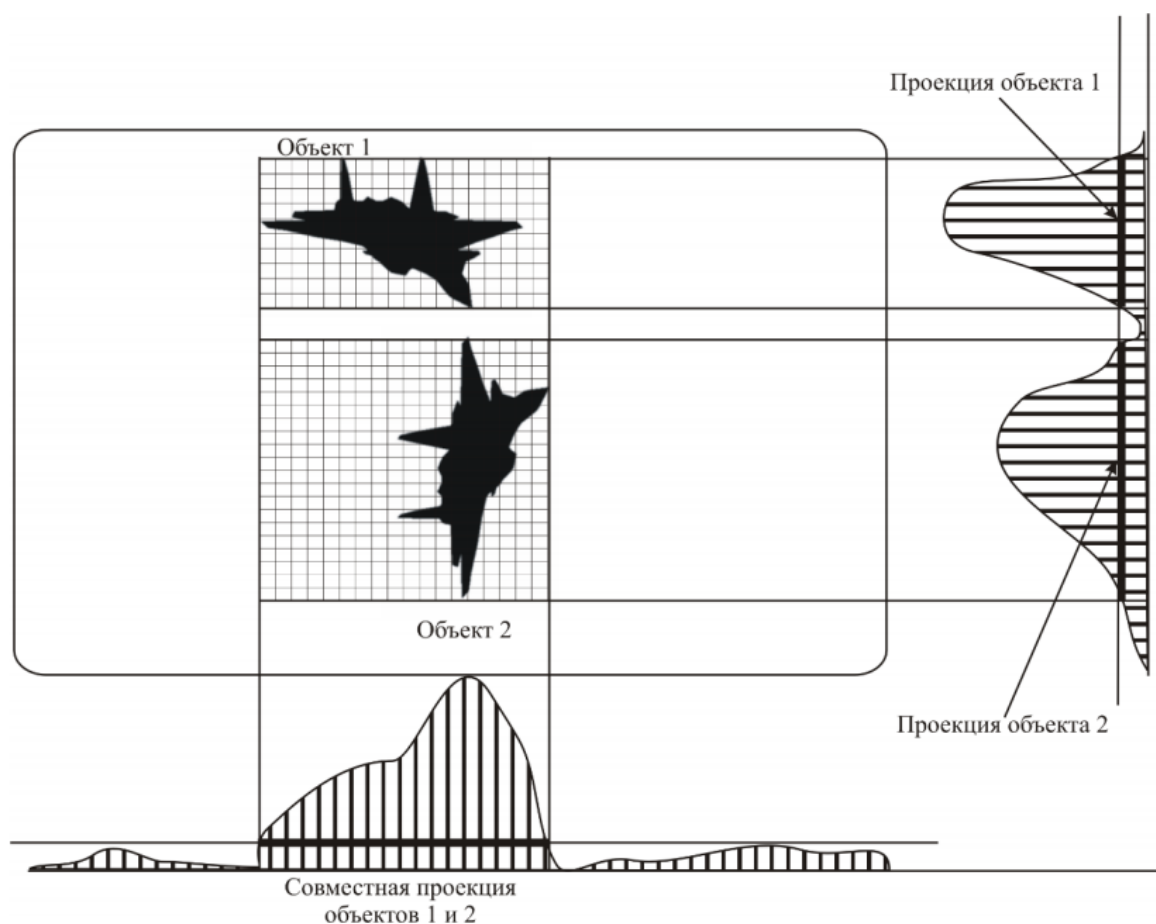


Рис. 4.4. Стробирование объектов

Если проекции объектов перекрывают друг друга, то либо не удастся установить центр тяжести объекта (при перекрытии по одному направлению, либо будут выделены сразу два объекта (при перекрытии по двум направлениям), что будет являться ошибкой сегментации.

Для того чтобы разрешить эту ситуацию возможно использование дополнительной координатной оси (например, оси под углом 45 градусов к координатным).

К ограничениям алгоритма нужно отнести тот факт, что с его помощью локализуется место в кадре, где зафиксировано движение. Однако о направлении и скорости движения посредством данного алгоритма судить нельзя. В результате даже если пара объектов движется в противоположные стороны с различными скоростями, но их проекции перекрываются, разделить данные объекты будет невозможно. Также без данных о направлении и скорости в общем случае невозможно реализовать трекинг объектов.

Следует отметить, что в данном случае рассмотрен алгоритм сегментации движущихся объектов при отсутствии в сцене глобального движения. То есть предполагается, что камера неподвижна и фон статичен. В случае, если камера перемещается, то применение описанного алгоритма возможно только после предварительной компенсации глобального движения.

Задание на моделирование

Разработать программу, реализующую алгоритм сегментации движущегося объекта с помощью оценки энергии движения. Исходными данными являются два кадра видеопоследовательности со статичным фоном, на которых присутствует перемещающийся объект.

В программе должны быть реализованы следующие процедуры, с выводом результата на экран:

- загрузка пары кадров,
- получение препарата межкадровой разности,
- получение контурного препарата с помощью одного из алгоритмов, указанных преподавателем: Собеля, разницы гауссиан, Кэнни, Робертса.
- объединение препаратов по логическому «И»
- обработка сформированного изображения с помощью морфологических фильтров (фильтрация и заливка),
- наложение строба на объект.

Контрольные вопросы

1. Каким образом формируют препарат межкадровой разности изображений?
2. Перечислите основные способы получения контурного препарата.
3. Раскройте принцип объединения изображений в соответствии с операцией логического «И»?
4. Поясните основные достоинства алгоритма сегментации объектов на основе энергии движения.
5. Поясните основные ограничения алгоритма сегментации на основе энергии движения.