

УДК 621.856.8

М.О. Абрамов, М.Ю. Катаев

Влияние методов предобработки на восстановление фигуры движущегося человека из потока изображений

Приводится сравнительное описание нескольких алгоритмов предварительного этапа обработки изображений в задаче выделения движущегося человека в потоке изображений. Приводятся результаты сравнительного анализа алгоритмов применительно к выделению фигуры человека. На основе проведенного анализа приводится критический анализ возможности алгоритмов.

Ключевые слова: алгоритмы предварительной обработки изображений, анализ, фигура человека, распознавание движения.

Актуальность задачи обработки статических и динамических изображений обусловлена применением в различных предметных областях, например, в таких областях науки, как астрономия, физика, медицина, а также промышленности, контроле, слежении и др. Процесс обработки изображений определяется целой последовательностью операций [1–7], в результате которых получается решение. Эта последовательность связана с этапами предварительной обработки, тематической обработки и анализа полученной информации. Результаты анализа зависят от качества работы соответствующих алгоритмов, которые используются в процессе предыдущих этапов. Разработано множество алгоритмов за годы развития методов на различных этапах обработки изображений. Отметим, что общего правила выбора того или иного набора алгоритмов пока не существует и все определяется сложностью решаемой задачи, размером и качеством получаемых изображений. Поэтому нами проводится анализ различных алгоритмов на этапе предварительной обработки в задаче выделения фигуры движущегося человека в потоке изображений.

Выделение движущегося человека в потоке изображений, полученных с помощью типовых видеокамер, является важным при решении задач спорта, контроля, слежения, медицины, промышленности и др. Особенность задачи в том, что при выделении исследуемого объекта в потоке изображений происходит изменение условий освещения, размеров, контрастности и других параметров. В таких условиях трудно подобрать один метод или набор алгоритмов, которые одинаково эффективно позволяли бы выделять объект исследования без искажений.

Основой для данной статьи является работа [8], в которой изложены первичные наработки в области выделения фигуры на изображении. Опыт массовой обработки видеофрагментов движения человека в потоке изображений способствовал появлению темы, которая отражена в данной статье, где приводится анализ алгоритмов предварительной обработки изображений в задаче выделения фигуры движущегося человека. Точность решения этой задачи определяет результаты анализа.

Постановка задачи. Анализ литературы [1–7] показал, что существует всего два подхода к выделению движущейся фигуры человека в потоке изображений: 1) выделение контура фигуры и 2) выделение площади фигуры. Оба этих подхода взаимосвязаны в одной из фаз обработки изображения, когда на основе алгоритмов первого подхода на конечном этапе можно оценить площадь фигуры. Так или иначе алгоритмы этих подходов в основе своей существенно отличаются. Качество работы алгоритмов обоих подходов будет зависеть от этапа предварительной обработки изображения. В работе нами используются алгоритмы выделения фигуры человека на изображении второго типа.

Рассмотрим математическую модель изображения, которая представляет собой матрицу структуры $\{RGB\}$ вида

$$I_{i,j} = m_{i,j} \cdot O_{i,j} + (1 - m_{i,j}) \cdot f_{i,j}, \quad (1)$$

где $I_{i,j}, i=1, \dots, N, j=1, \dots, M$ – наблюдаемое изображение, $O_{i,j}$ – часть изображения, занимаемая объектом; $m_{i,j}$ – бинарная маска области занимаемой объектом (1 соответствует части изображения, занимаемой объектом, 0 – части изображения, относящейся к фону); $f_{i,j}$ – значения пикселей фона.

Задача этапа выделения фигуры связана с нахождением бинарной маски m . Одним из вариантов решения задачи является работа с разностью последующих кадров, на которых, по сути, пиксели, соответствующие фону, должны быть удалены. Однако за время между кадрами (для типичных камер 0,02–0,04 с) происходят достаточные изменения освещенности, возникает эффект движущихся теней, появляются новые объекты фона и др., которые не позволяют однозначно восстановить фигуру человека. Также на изображении присутствуют шумы различного класса, которые также способствуют усложнению.

В качестве основного алгоритма нахождения маски мы используем адаптивную гауссовскую смешанную модель [9, 10]. Это самообучающийся алгоритм, основанный на определении параметров гауссовского распределения, среднего для каждого пикселя потока изображений.

Сглаживающие фильтры. Для удаления шумов в цифровой обработке используют различного рода фильтры. Они могут использоваться как на этапе предобработки, так и постобработки изображений. Шумы бывают нескольких типов [1–7]: соль и перец, белый шум – случайные белые пиксели; гауссов шум – колебания яркости распределяются по нормальному закону и др. Поэтому фильтры, как правило, делятся по сложности и способности удаления того или иного типа шумов. Мы решили включить в обработку четыре наиболее распространённых фильтра: линейный; медианный; гауссовский и двусторонний.

Наиболее простым является линейный фильтр. Каждая точка кадра последовательно обрабатывается. Значение вычисляется с помощью следующей формулы [11]:

$$In_{i,j} = \sum_{k,l} I_{i+k,j+l} \cdot H_{k,l}, \quad (2)$$

где In – результирующее значение пикселя нового изображения (n); I – исходное значение пикселя; k, l – индексы окна сглаживания ($k=1, \dots, k_{\text{width}}, l=1, \dots, k_{\text{height}}$); H – ядро, представляющее собой коэффициенты линейного фильтра, представляется в виде:

$$H = \frac{1}{k_{\text{width}} \cdot k_{\text{height}}} \begin{vmatrix} 1 & \dots & 1 \\ \dots & \dots & \dots \\ 1 & \dots & 1 \end{vmatrix}.$$

Элементы матрицы ядра сглаживающего фильтра могут меняться в зависимости от типа применяемого сглаживания.

Такой фильтр создаёт размытие изображения и приводит к сглаживанию резких перепадов яркости. Эти процедуры являются оптимальными при гауссовском распределении помех. В случае импульсных помех, разбросанных по кадру, эффективность этого метода резко падает. Удачным решением возникающей проблемы является применение медианного фильтра. Также как и в линейном фильтре, пиксели изображения обрабатываются последовательно, а для вычисления значения используется некоторая окрестность, представленная, как правило, в виде креста или квадрата. Её размеры определяются в зависимости от задачи и характера изображения. Точки, попавшие в пределы окрестности, образуют рабочую выборку текущего шага. Если упорядочить выборку по возрастанию, то её медианой будет тот элемент выборки, который займёт центральное положение. Фильтр успешно обрабатывает импульсные выбросы яркости, однако если импульсная помеха не является точечной, а покрывает некоторую локальную область, она может быть исключена лишь в случае, если размер выборки превышает размер этой области, поэтому данный фильтр также сталкивается с некоторыми трудностями.

Другим способом шумоподавления является применение гауссовского фильтра с ядром [12]:

$$H_{ij} = \frac{1}{2\pi\sigma^2} \cdot \exp\left\{-\frac{i^2 + j^2}{2\sigma^2}\right\}, \quad (3)$$

где σ – дисперсия; i, j – координаты точки.

В этом подходе пиксели в некотором выбранном по размеру окне ($k_{\text{width}}, k_{\text{height}}$) смешиваются по закону, заданному функцией Гаусса. Влияние пикселей друг на друга при гауссовской фильтрации обратно пропорционально квадрату расстояния между ними.

Представленные выше фильтры хорошо справляются с очищением шума на основной площади изображения и плохо сглаживают края (краевой эффект). Во избежание этого стоит использовать так называемый – двусторонний фильтр. По аналогии с гауссовским фильтром он также учитывает веса

соседних пикселей. Эти веса состоят из двух частей, первая рассчитывается так же, как и в гауссовском фильтре. Вторая представляет собой разность в интенсивности между соседним и вычисляемым пикселем. Модель двустороннего фильтра записывается следующим образом [13]:

$$In(x) = G^{-1}(x) \cdot \iint I(x)q(x',x)s(I(x'),I(x))dx' \quad (4)$$

с множителем нормировки $G(x)$ вида:

$$G(x) = \iint q(x',x)s(I(x'),I(x))dx',$$

где $x = (i, j)$ – пиксель изображения с определенными координатами; $In(x)$ – значение пикселя x в выходном изображении; $I(x)$ – значение пикселя x во входном изображении; $q(x',x)$ – геометрическая близость, расстояние между центром окрестности x и соседней точкой x' , $s(I(x'), I(x))$ – близость значений пикселей, разброс значений пикселей между центром окрестности x и соседней точкой x' .

Фильтры цветокоррекции. Цветокоррекция в отличие от фильтров преобразует значения цветов пикселей, в зависимости лишь от его прежнего значения:

$$In(x) = F(I(i, j)),$$

здесь F – функция преобразования цветовой палитры.

Основная причина, по которой приходится выполнять коррекцию цвета, заключается в том, что человеческий глаз, в отличие от камеры, умеет адаптироваться к силе и спектральными характеристикам таким образом, что сохраняется восприятие цвета предметов в большинстве случаев, независимо от спектрального состава освещения. При просмотре снятых камерой изображений в других условиях мы увидим, что они отличаются от того, что мы видели, когда записывали. Также цветокоррекция применяется в следующих случаях: недостаточный или избыточный контраст изображения, съемка при плохих погодных условиях (туман, дождь и т.д.), плохое освещение и т.д.

Основные алгоритмы этапа предварительной обработки изображений связаны с применением методов приведения изображений к одинаковым условиям по освещению, удалению влияния шума на изображении, коррекции контраста и цветности, а также изменениям цветового пространства. Далее будут кратко рассмотрены основные подходы, которые нами применяются в работе.

На первом этапе предварительной обработки, все последовательности изображений приводятся к балансу белого цвета по известной методике «Серый мир». Далее нами применяются сглаживающие фильтры в том или ином цветовом пространстве.

Преобразование «Серый мир» изменяет цвета пикселей таким образом, что сумма всех цветов результирующего изображения даёт серый цвет. Модель преобразований выглядит следующим образом:

$$In(x) = I(x) \cdot \frac{Avg}{\overline{I(x)}}, \quad (5)$$

где $\overline{I(x)}$ – среднее значение канала исходного изображения; Avg – среднее значение всех каналов исходного изображения. Каналы (RGB или HSV и др.) могут изменяться в зависимости от модели цветового пространства изображения.

При выполнении процедуры контрастирования изображения строится гистограмма интенсивности пикселей, после чего она преобразуется с помощью кумулятивной функции распределения [10]. Значения пикселей результирующего изображения извлекаются из полученной гистограммы и поступают на следующее преобразование

Помимо перечисленных выше процедур может возникнуть необходимость изменения цветового пространства. Это обусловлено тем, что, к примеру, в цветовом пространстве HSV тень человека распознаётся лучше, чем в цветовом пространстве RGB , это связано с тем, что тень, как правило, обладает высоким значением тона – H , высокой насыщенностью – S и низким значением – V . Согласно [14] мы решили включить три наиболее подходящих цветовых пространства:

- 1) RGB (red green blue) – цветовая модель, согласно которой пиксель экрана получается путём наложения различных величин трех цветов на черный.
- 2) HSV (Hue, Saturation, Brightness) – пиксель получается путём указания тона, насыщенности и яркости.
- 3) YUV – яркость и две цветоразностных.

Описание процесса моделирования. Входными параметрами является массив, состоящий из 20 изображений разрешением 640×480 пикселей, отображающий движение человека перпендикулярно камере справа налево на расстоянии приблизительно три метра. Параметром оценки качества

работы методов выступает площадь (S) найденной фигуры человека. Площадь вычисляется у наибольшего контура из найденных на результирующих бинарных изображениях, отображающих распознанное смешанным гауссовским адаптивным методом движение выделенными «белыми» пикселями. После чего определяется её среднее значение, которое и является оценкой (\bar{S}).

Алгоритм программы последовательно обрабатывает изображения в соответствии с заданным циклом. Элементами цикла являются рассмотренные методы предобработки. Цикл отображается на рис. 1.

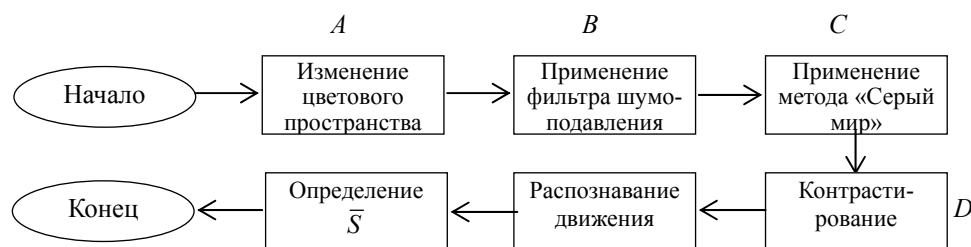


Рис. 1. Схема работы программы

Согласно схеме работы программы, блоки A , B , C , D (см. рис. 1) могут изменяться в зависимости от выбора цепочки методов предобработки. Каждый блок в процессе обработки, может включать только один вариант из методов, представленных в табл. 1.

Таблица 1

Номера параметров моделирования			
Цветовое пространство	Фильтр	«Серый мир»	Контрастирование
RGB = 1	Гауссовский = 1	Включен = 1	Включено = 1
HSV = 2	Медианный = 2		
IUV = 3	Двусторонний = 3	Выключен = 0	Выключено = 0
	Линейный = 4		
A	B	C	D

Таким образом, согласно таблице параметров (табл. 1) и схеме работы программы (рис. 1) цикл предобработки, включающий в себя цветовую модель HSV , двусторонний фильтр и контрастирование, будет выглядеть в соответствии с последовательностью блоков $A-B-C-D$, например: 2-3-0-1.

В соответствии с описанными выше параметрами было проведено исследование влияния рассмотренных методов предобработки на восстановление фигуры движущегося человека на изображении.

Результаты моделирования. Учитывая число сочетаний методов, было произведено максимальное число опытов, возможное при данных входных параметрах, а именно $3 \times 4 \times 2 \times 2 = 48$ процедур моделирования со всеми возможными комбинациями методов предобработки. После каждого цикла обработки найдена средняя площадь всех обработанных кадров \bar{S} . Также было проведено ручное выделение движущегося объекта на входных изображениях, средняя площадь составила $\bar{S}_0 = 30830,52$ пикселей, данное значение можно считать эталонным с минимальной погрешностью. В последней колонке табл. 1 обозначено процентное отклонение от эталонного значения у того или иного метода, рассчитываемое по формуле

$$\Delta \bar{S} = 100 \cdot \left(\frac{\bar{S}}{\bar{S}_0} - 1 \right).$$

Из табл. 2 видно, что в основном восстанавливаемая фигура человека по площади меньше исходной (знак минус), при этом большинство вариантов сочетаний методов дают погрешность от 30–50%. Серым выделено всего три случая, когда погрешность составила не более 15% (номера 24, 26 и 28). Также наблюдаем, что наибольшие скачки возникают при использовании цветового пространства YUV и HSV . В то время как пространство RGB дает наиболее стабильные результаты небольшого отклонения, при использовании цветового пространства YUV резкие отклонения наблюдаются в местах, где отсутствует метод контрастирования, это обусловлено особенностями цветового пространства. Лучшие результаты получились при использовании каналов HSV , медианного и двустороннего фильтров при применении контрастирования.

Таблица 2

Сравнение результатов вычисления площади фигуры

<i>N</i>	Методы	ΔS	<i>N</i>	Методы	ΔS	<i>N</i>	Методы	ΔS
1	1-1-0-0	-35,65	17	2-1-0-0	-64,08	33	3-1-0-0	-56,78
2	1-1-0-1	-28,66	18	2-1-0-1	-45,72	34	3-1-0-1	-34,54
3	1-1-1-0	-35,59	19	2-1-1-0	-37,24	35	3-1-1-0	-52,39
4	1-1-1-1	-28,51	20	2-1-1-1	-28,93	36	3-1-1-1	-34,89
5	1-2-0-0	-34,15	21	2-2-0-0	-55,57	37	3-2-0-0	-55,91
6	1-2-0-1	-27,86	22	2-2-0-1	-72,54	38	3-2-0-1	-33,27
7	1-2-1-0	-34,21	23	2-2-1-0	-28,43	39	3-2-1-0	-51,87
8	1-2-1-1	-28,51	24	2-2-1-1	-15,15	40	3-2-1-1	-33,18
9	1-3-0-0	-33,46	25	2-3-0-0	-41,84	41	3-3-0-0	-55,48
10	1-3-0-1	-28,35	26	2-3-0-1	-7,74	42	3-3-0-1	-32,64
11	1-3-1-0	-33,49	27	2-3-1-0	-21,30	43	3-3-1-0	-52,31
12	1-3-1-1	-28,26	28	2-3-1-1	-11,91	44	3-3-1-1	-32,53
13	1-4-0-0	-35,94	29	2-4-0-0	-71,31	45	3-4-0-0	-57,29
14	1-4-0-1	-27,43	30	2-4-0-1	-76,39	46	3-4-0-1	-34,28
15	1-4-1-0	-35,86	31	2-4-1-0	-42,18	47	3-4-1-0	-52,64
16	1-4-1-1	-27,34	32	2-4-1-1	-36,41	48	3-4-1-1	-34,31

Как видно из табл. 2, найденные площади фигуры занижены для всех методов, а также, судя по таблице результатов (табл. 1), значения меняются в диапазоне (-77%; -7%). Такой большой диапазон и тенденция занижения, в первую очередь, обусловлены тем, что после каждой обработки происходит удаление всех наименьших контуров, тем самым очищая результат от помех. Любой найденный каким-либо методом нецелостный силуэт в таком случае обрезается. Для избегания подобной ситуации применяются морфологические операции сужения и расширения. На следующем этапе нами будет исследовано точное влияние данных операций на результат, но в данной статье этот фактор не учитывается.

На рис. 2 приведены средние процентные отклонения найденной площади в зависимости от методов предобработки изображений.

а	Цветовое пространство	
	RGB = − 31,45	
	HSV = − 41,05	
	IUV = − 44,02	

б	Фильтр	
	Гауссовский = − 44,65	
	Медианный = − 39,22	
	Двусторонний = − 31,61	

в	Серый мир		
	Контрастирование	1	0
	1	−28,33	−37,45
	0	−39,79	−49,79

Рис. 2. Средние процентные отклонения найденной площади в зависимости от методов предобработки

Согласно рис. 2, *a*, который отображает средние отклонения площади от эталонной по каждому из методов предобработки, изменение цветового пространства создает колебания 31–44% с отличием друг от друга до 13%. Различия в отклонении площади при замене фильтров сглаживания (см. рис. 2, *б*) также не превышает 31–44% при отклонении до 13%. Операции контрастирования и «Серый мир» (рис. 2, *в*) при различном комбинировании изменяют результат в диапазоне 28–50% при отклонении между собой до 22%. Следовательно, можно отметить, что степень их воздействия на результат является максимальной, и необходимо тщательно относиться к этой процедуре предварительной обработки изображений, подбирая сочетание лучших алгоритмов.

Заключение. Стоит отметить, что данное исследование не является всеохватывающим в том смысле, что все исследования проводятся лишь на одной последовательности изображений, без изменения различных фоновых составляющих и пр. Поэтому их воздействие на методы предобработки не описывается. Однако оно даёт возможность увидеть картину в целом, что в дальнейшем может способствовать более глубокому изучению влияния методов предобработки на процессы распознавания движущихся объектов на изображениях. Главное, что необходимо отметить, – влиянием методов предварительной обработки изображений нельзя пренебрегать. В зависимости от типа

сцены на изображении и условий освещенности необходимо проводить подобные модельные расчеты, которые приведены в данной статье.

Данная статья выполнена в рамках проекта Российского гуманитарного научного фонда (РГНФ) N:14-16-70008 а(р).

Литература

1. Яне Б. Цифровая обработка изображений / Б. Яне. – М.: Техносфера, 2007. – 584 с.
2. Прэтт У. Цифровая обработка изображений: в 2 т. / У. Прэтт. – М.: Мир, 1982.
3. Гонсалес Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс. – М.: Техносфера, 2005. – 1072 с.
4. Абламейко С.В. Обработка изображений: технология, методы, применение: учеб. пособие / С.В. Абламейко, Д.М. Лагуновский. – М.: Амафед, 2000. – 304 с.
5. Методы компьютерной обработки изображений / ред. В.А. Сойфер. – М.: Физматлит, 2001. – 784 с.
6. Цифровая обработка изображений в информационных системах: учеб. пособие / И.С. Грузман, В.С. Киричук и др. – Новосибирск: НГТУ, 2002. – 352 с.
7. Лукьяница А.А. Цифровая обработка видеоизображений / А.А. Лукьяница, А.Г. Шишкин. – М.: Ай-Эс-Си Пресс, 2009. – 518 с.
8. Определение и анализ двигательной активности постинсультного пациента из потока изображений / М.Ю. Катаев, С.Г. Катаев, Н.Г. Катаева, В.А. Чистякова // Информатика и системы управления. Медицинская информатика. – 2012. – № 4 (34). – С. 43–50.
9. Kaew Trakulpong P. An improved adaptive background mixture model for real-time tracking with shadow detection / P. Trakulpong Kaew, R. Bowden // Kluwer Academic Publishers. – 2002. – P. 135–144.
10. Вентцель Е.С. Теория вероятностей: учеб. для вузов / Е.С. Вентцель. – М.: Высш. шк., 1999. – 576 с.
11. Richard S. Computer Vision: Algorithms and Applications / S. Richard. – London: Springer, 2011. – 655 p.
12. Forsyth D.A. Computer Vision: A Modern Approach: 2nd Edition / D.A. Forsyth, J. Ponce. – New Jersey: Prentice Hall, 2011. – 792 p.
13. Tomasi C. Bilateral filtering for gray and color images / C. Tomasi, R. Manduchi // Computer Vision, Sixth International Conference. – Washington, DC, USA, 1998. – P. 839–846.
14. Kumar P. A Comparative Study of Different Color Spaces for Foreground and Shadow Detection for Traffic Monitoring System / P. Kumar, K. Sengupat, A. Lee // Journal of Intelligent Transportation Systems: Technology, Planning, and Operations. – 2002. – No 7. – P. 100–105.

Абрамов Максим Олегович

Аспирант каф. оптико-электронных систем и дистанционного зондирования
Национального исследовательского Томского государственного университета
Тел.: 8 (913) 117-78-30
Эл. почта: maxim_amo@mail.ru

Катаев Михаил Юрьевич

Д-р техн. наук, профессор каф. автоматизированных систем управления (АСУ) ТУСУРа,
профессор Юргинского технологического института (филиала)
Национального исследовательского Томского политехнического университета
Тел.: 8-960-975-27-85, (382-2) 70-15-36
Эл. почта: kataev.m@sibmail.com

Abramov M.O., Kataev M.Y.

The effect of preprocessing methods on figure of the moving human retrieving from an images sequence

The article contains a comparative description of several algorithms of image preprocessing for issue of detection moving of human in the stream of images. Provides results of comparative analysis of algorithms. On the basis of the results provides a critical analysis of the capabilities of algorithms.

Keywords: pre-processing algorithms of images, figure of human, analysis, motion detection.