

Алгоритм динамической сегментации пары последовательных кадров

С.Е. Ваганов¹

¹ Ивановский государственный университет, Иваново, Россия

Аннотация

Разработан алгоритм динамической сегментации пары кадров видео. Проведен сравнительный анализ качества сегментации для случаев поиска сдвигов и аффинных межкадровых преобразований для сегментов. Также произведено сравнение с некоторыми методами статической сегментации.

Ключевые слова: сегментация, изображение, видео, аффинное преобразование, оптический поток.

Цитирование: Ваганов, С.Е. Алгоритм динамической сегментации пары последовательных кадров / С.Е. Ваганов // Компьютерная оптика. – 2019. – Т. 43, № 1. – С. 83-89. – DOI: 10.18287/2412-6179-2019-43-1-83-89.

Введение

Сегментация видео (motion segmentation) находит свое применение при решении многих задач компьютерного зрения, в частности: слежение за объектами и их классификация, анализ видео, выделение фона, уточнение оптического потока для интерполяции (экстраполяции) кадров и многие другие. Основная цель, стоящая перед алгоритмами сегментации видео [1–4], заключается в разделении всех точек некоторого кадра на подвижные объекты и фон [5].

В зависимости от решаемой задачи, понятие сегмента (объекта) может трактоваться различным образом. В данной работе мы строим сегментацию оптического потока по паре последовательных кадров видео. Сегментом называем совокупность точек первого кадра, имеющих схожее межкадровое движение, определяемое оптическим потоком.

Первая из основных проблем сегментации видео связана с перекрытиями, появлениями новых сегментов и исчезновением имеющихся. В этом случае поиск межкадрового отображения (преобразования) сегмента является нетривиальной задачей.

Если вектора сдвигов в оптическом потоке достаточно точно описывают модель межкадрового движения отдельных точек, то задача сегментации упрощается. В противном случае сегменты будут найдены недостаточно точно. Однако даже в этом случае построенная сегментация позволяет улучшить оптический поток. Таким образом, задачу динамической сегментации нужно решать в паре с задачей уточнения оптического потока, чему и посвящена настоящая работа.

Оптический поток позволяет решать большое число задач компьютерного зрения и цифровой обработки видео. Самыми распространенными можно назвать задачи временной интерполяции кадров и сопоставления стереопары. Несмотря на то, что на сегодняшний день существует много способов построения оптического потока [6], в том числе и основанных на сегментации, число исследований в данной области продолжает расти, что свидетельствует об актуальности данной проблемы.

В работе [5] предлагается разделить все алгоритмы сегментации видео на несколько категорий:

- 1) разностные методы (*Image difference*) – методы пороговой сегментации попиксельной разности пары кадров;
- 2) основанные на кластеризации особых точек в пространстве признаков (*Manifold clustering*);
- 3) основанные на статистических фреймворках (EM, PM, MAP и др.);
- 4) основанные на вейвлет-анализе;
- 5) сегментирующие оптический поток;
- 6) основанные на сегментации на слои (*Layers*) – каждый слой содержит набор точек со схожей моделью межкадрового движения.

Согласно данной классификации, рассматриваемый в работе подход включает в себя некоторые основные идеи методов из классов 5–6.

В работе [4] используются подходы, близкие к сегментации, однако авторы не требуют, чтобы движение объекта описывалось одной формулой. Поэтому сегментация получается иной.

Целью работы являлась разработка алгоритма динамической сегментации пары последовательных кадров видео, формирующего сегменты, движение которых хорошо описывается геометрическими межкадровыми преобразованиями (например, посредством сдвигов или аффинных преобразований). Предложенный подход включает в себя процедуры склейки сегментов и уточнения их границ, которые не зависят от типа преобразований сегментов. Для вычисления межкадровых преобразований используется оптический поток.

Кроме того, проведено сравнение предложенного алгоритма с использованием различных межкадровых моделей движений сегментов, а также с некоторыми подходами к сегментации отдельных изображений.

В п. 1 фиксируются некоторые обозначения, используемые в дальнейшей работе. В п. 2 приводится описание подхода к нахождению аппроксимации аффинного преобразования сегмента по известному оптическому потоку. В п. 3 – описание разработанного алгоритма сегментации. В п. 4 – достигнутые результаты, в частности показывающие, что предложенный алгоритм динамической сегментации даёт существенно лучшее качество по сравнению со статическими методами.

1. Основные понятия

Пусть f – изображение размера $m \times n$ (матрица целых чисел от 0 до 255). Рассматривая f как таблично заданную функцию с областью определения $U = \{u_i \in R^2, i = 0, m \cdot n - 1\}$, продолжим её на всю действительную плоскость с помощью некоторой интерполяционной формулы. В точках, расположенных за границами изображения, значение функции считаем равным значению на границе.

В дальнейшем через f и g будем обозначать пару последовательных кадров видео.

Межкадровое движение может быть описано преобразованием $T(x, y)$, представляющим собой отображение точки $u = (x, y) \in U$ с первого кадра на второй. В работе рассматриваются сдвиги и аффинные преобразования:

$$T(x, y) = (x + a_0, y + a_1), \tag{1}$$

$$T(x, y) = (a_0 x + a_1 y + a_2, a_3 x + a_4 y + a_5), \tag{2}$$

где x, y – координаты точки на изображении, а $a_i \in R$.

В классических алгоритмах поиска межкадрового движения [7, 8] используется (1).

Разбиение $U = S_0 \cup S_1 \cup \dots \cup S_{N-1}$ на непустые подмножества S_i (сегменты) будем называть сегментацией.

На практике к сегментам обычно предъявляются некоторые дополнительные требования типа связности, гладкости, однородности цвета (яркости) и т.п.

Пусть $S \subset U$ – сегмент, а T – отображение точек S с кадра f на кадр g , тогда качество пары $\{S, T\}$ будем оценивать посредством среднеквадратичного отклонения:

$$Q_S(S, T) = \sqrt{\frac{1}{|S|} \cdot \sum_{u_i \in S} (f(u_i) - g(T(u_i)))^2}, \tag{3}$$

где $|S|$ – количество точек сегмента.

Динамической сегментацией $\{S_i, T_i\}$ будем называть разбиение $U = S_0 \cup S_1 \cup \dots \cup S_{N-1}$, дополненное набором межкадровых преобразований T_i для каждого S_i , таким, что $Q_S(S_i, T_i) \rightarrow \min$. Неформально говоря, межкадровое движение точек S_i должно хорошо описываться преобразованием T_i .

Аналогичным образом будем производить оценку качества сегментации [9]:

$$Q = \sqrt{\frac{1}{m \cdot n} \cdot \sum_{i=0}^{N-1} \sum_{u_j \in S_i} (f(u_j) - g(T_i(u_j)))^2}, \tag{4}$$

где N – количество сегментов.

В пределе, когда каждый сегмент состоит из одной точки, величину Q можно сделать очень малой, однако количество сегментов при этом будет чрезвычайно велико (число пикселей изображения). Мы хотим минимизировать величину Q , не увеличивая количество сегментов.

Из произвольной статической сегментации первого кадра пары можно получить динамическую, если для каждого сегмента S_i найти преобразование T_i , минимизирующее величину (3). Такой подход позволяет

сравнивать статические подходы к сегментации с динамическими.

2. Нахождение межкадровых преобразований

Векторное поле, описывающее движение каждой точки с кадра f на кадр g , будем называть оптическим потоком [10].

Оптический поток находит свое применение во многих задачах компьютерного зрения и обработки видео, в частности: обнаружение и отслеживание положения объекта, реконструкция 3D-сцены по набору изображений, сегментация и обнаружение границ объектов, нахождение фона, вычисление дифференциальных кадров при сжатии видеoinформации [11] и т.д.

В настоящей работе для нахождения оптического потока использовался метод [12], так как он показал более устойчивый с точки зрения величины вариации поля векторов результат, чем классический алгоритм Лукаса – Канаде [7].

Задача нахождения межкадровых преобразований T для сегмента S сводится к задаче минимизации следующего функционала:

$$\sum_{u_i \in S} (f(u_i) - g(T(u_i)))^2 \rightarrow \min. \tag{5}$$

Другими словами, для области S требуется найти преобразование (в нашем случае вида (1) или (2)) на паре кадров f и g такое, чтобы сумма квадратов отклонений значений функций f и g в данных точках была минимальна.

При найденном оптическом потоке нахождение преобразований для сегмента значительно упрощается. Аппроксимация вектора сдвига для S вычисляется посредством усреднения сдвигов всех точек сегмента S по найденному оптическому потоку. Для нахождения аффинных преобразований вместо минимизации функционала (5) будем минимизировать отклонение искомого преобразования (2) от имеющегося потока.

Пусть $V(x, y) = (V_x, V_y)$ – матрица векторов сдвигов точек с кадра f на кадр g (оптический поток). Тогда для нахождения аппроксимации аффинного межкадрового преобразования (2) сегмента S необходимо решить две системы линейных уравнений размера 3×3 , а именно:

$$A \cdot X = b_x, \tag{6}$$

$$A \cdot Y = b_y, \tag{7}$$

где A , b_x и b_y определяются следующим образом:

$$A = \sum_{(x,y) \in S} \begin{pmatrix} x^2 & x \cdot y & x \\ x \cdot y & y^2 & y \\ x & y & 1 \end{pmatrix}, \tag{8}$$

$$b_x = \sum_{(x,y) \in S} ((x^2 \ x \cdot y \ x)^T + V_x(x, y) \cdot (x \ y \ 1)^T), \tag{9}$$

$$b_y = \sum_{(x,y) \in S} ((x \cdot y \ y^2 \ y)^T + V_y(x, y) \cdot (x \ y \ 1)^T). \tag{10}$$

Таким образом, решив системы (6) и (7), получим пару векторов, элементы которых представляют со-

бой значения весовых коэффициентов искомой аппроксимации аффинного межкадрового преобразования (2) сегмента S с кадра f на кадр g :

$$X = (a_0 \ a_1 \ a_2)^T = A^{-1} \cdot b_x, \quad (11)$$

$$Y = (a_3 \ a_4 \ a_5)^T = A^{-1} \cdot b_y. \quad (12)$$

Замечание 1. Решение (6) и (7) существует и единственно тогда и только тогда, когда точки сегмента не лежат на одной прямой.

Замечание 2. Точность найденного преобразования будет зависеть от точности оптического потока и размера сегмента.

Метод нахождения аффинных преобразований посредством минимизации (5), описанный в работе [13], позволяет получить более точную аппроксимацию, но имеет существенно меньшую производительность, чем (11), (12). Поэтому вычисление аппроксимации аффинного преобразования посредством приведенного выше подхода оказалось предпочтительным.

3. Алгоритм динамической сегментации

На входе алгоритм получает пару кадров (f и g) и некоторую начальную сегментацию кадра f .

Алгоритм динамической сегментации включает в себя выполнение следующих шагов:

- 1) нахождение межкадровых преобразований с кадра f на кадр g для всех S_i ;
- 2) вычисление текущего качества сегментации Q_0 по формуле (4);
- 3) склейка соседних сегментов с учетом стоимости объединения;
- 4) уточнение границ сегментов;
- 5) вычисление качества Q_1 по формуле (4);
- 6) повторение шагов 2–5, пока осуществляется склейка сегментов или $|Q_0 - Q_1| > \epsilon$.

Склейка соседних сегментов с учётом стоимости объединения

Пусть S_i и S_j – пара смежных сегментов с межкадровыми преобразованиями T_i и T_j . Тогда стоимостью объединения S_i с S_j будем называть разность:

$$Cost(S_i, T_i, T_j) = Q_s(S_i, T_j) - Q_s(S_i, T_i). \quad (13)$$

Значение стоимости объединения (13) позволяет определить, на какую величину изменится качество сегмента при изменении его геометрического преобразования. Если значение $Cost(S_i, T_i, T_j) > 0$, то качество ухудшилось.

Алгоритм склейки:

1) Повторяем, пока существуют сегменты с $\text{Cos } i(S_i, T_i, T_j) < \text{Theshold}$: к сегменту S_i присоединяем соседний сегмент S_j , имеющий наименьшее значение $\text{Cos } i(S_i, T_i, T_j)$. Пересчитываем преобразование T_i для сегмента $S_i = S_i \cup S_j$.

2) Если слияния сегментов (в шаге 1) не осуществлялись или $\text{Theshold} \geq \max \text{Theshold}$, то алгоритм склейки завершает свою работу.

3) Увеличиваем максимальную стоимость объединения Theshold на фиксированную величину α и переходим к шагу 1.

Уточнение границ сегментов

Точку $p \in S_i$ называем граничной, если хотя бы одна из 8 соседних точек не принадлежит S_i .

Пусть p – точка кадра f . Тогда качество $p \in S_i$ представимо в виде:

$$Q_{pnt}(p, T_i) = \sqrt{(f(p) - g(T_i(p)))^2}. \quad (14)$$

Алгоритм уточнения границ сегментов:

1) Если точка $p \in S_i$ граничит с сегментом S_j , то присоединяем p к S_j в случае уменьшения погрешности: $Q_{pnt}(p, T_j) < Q_{pnt}(p, T_i)$.

2) Если размер S_i меньше фиксированного порога MinSize , то его точки распределяются (с учетом качества (14)) между смежными сегментами.

3) Повторять шаги 1–2 для всех точек кадра f , пока имеет место модификация сегментов.

Данная процедура, состоящая из шагов 1–3, сходится за конечное число шагов, так как преобразования для сегментов не пересчитываются в процессе ее выполнения.

Замечание 3. Для уменьшения влияния шума на оценку качества точки с координатами (x, y) качество необходимо вычислять в виде средневзвешенной суммы (14) по окрестности фиксированного радиуса R (в настоящей работе $R = 1$).

Замечание 4. После пересчета преобразований сегментов или их склейки оптический поток модифицируется.

4. Оценка качества

Для тестирования работы алгоритма использовались следующие параметры:

α	maxThreshold	MinSize	ϵ
0,2	10	100	0,0005

Для уменьшения влияния шума на работу алгоритма все проверочные изображения [14] были размыты фильтром Гаусса с размером ядра 3×3 . При оценке качества полученной сегментации использовались исходные изображения.

Для оценки работы алгоритма в качестве начальной сегментации использовалось разбиение первого кадра на 400 прямоугольных равноразмерных сегментов.

Ниже приведены таблицы, в которых представлены оценки качества работы предложенного алгоритма по сравнению с методами статической сегментации, где Q и N – значения качества сегментации (4) и полученного числа сегментов для всех пар проверочных кадров [14].

Методы SQ (разбиение первого кадра на прямоугольные области равного размера), WAT (метод водораздела с автоматической расстановкой меток сегментов) и RS (сегментация посредством выделения непрерывных границ [15]) – некоторые алгоритмы статической сегментации (на N сегментов) первого кадра с последующим расчетом межкадровых преобразований полученных сегментов посредством использования метода, описанного в работе [16]. Методы DSS (Dynamic Shift Segmentation) и DAS (Dynamic Affine Segmentation) представляют собой реализацию

предложенного алгоритма динамической сегментации для случаев поиска сдвигов (1) и аффинных преобразований (2) сегментов.

Табл. 1. Оценка качества предложенного метода для случая поиска сдвигов (1) по сравнению со статической сегментацией

Метод \ Кадр	DSS		SQ		WAT		RS	
	N	Q	N	Q	N	Q	N	Q
000.bmp	18	3,36	16	6,15	19	5,98	19	5,94
004.bmp	7	8,64	9	10,14	8	11,00	8	10,18
300.bmp	21	8,56	25	14,39	22	15,73	22	15,44
500.bmp	10	6,03	9	11,18	11	10,76	11	10,81
504.bmp	16	3,10	16	4,26	17	4,58	17	4,32
600.bmp	11	5,14	9	6,54	12	6,57	12	7,51
604.bmp	33	5,13	36	8,35	34	9,03	34	9,34
616.bmp	11	6,42	9	10,32	12	10,42	12	8,45
618.bmp	17	4,04	16	6,13	18	6,15	18	6,08
619.bmp	17	7,14	16	8,47	18	9,14	18	11,30
621.bmp	16	4,45	16	8,53	17	8,42	17	9,08
Ср. Q	5,64		8,59		8,89		8,95	

Табл. 2. Оценка качества предложенного метода для случая поиска аффинных преобразований (2) по сравнению со статической сегментацией

Метод \ Кадр	DAS		SQ		WAT		RS	
	N	Q	N	Q	N	Q	N	Q
000.bmp	27	2,89	25	4,93	28	5,23	28	5,15
004.bmp	9	5,09	9	8,11	10	8,76	10	6,92
300.bmp	25	6,13	25	12,01	26	12,83	26	11,5
500.bmp	59	6,55	64	7,21	60	7,03	60	7,55
504.bmp	18	1,99	16	3,03	19	3,1	19	3,4
600.bmp	8	3,71	9	4,32	9	4,93	9	5,42
604.bmp	42	4,71	36	6,73	43	7,22	43	7,49
616.bmp	13	6,78	16	7,66	14	8,22	14	9,13
618.bmp	19	4,33	16	5,52	20	5,28	20	5,49
619.bmp	35	4,25	36	5,01	36	4,66	36	9,31
621.bmp	44	4,15	49	5,61	45	5,46	45	6,06
Ср. Q	4,6		6,38		6,61		7,04	

В последней строке представлена усредненная (по всем кадрам) оценка качества Q для соответствующих методов.

В случаях SQ , WAT и RS используется разбиение на соизмеримое с динамическим подходом (DSS и DAS) количество сегментов.

В среднем, метод, использующий поиск аффинных преобразований (2), показал лучшее качество по сравнению с алгоритмом, использующим поиск сдвигов (1).

По сравнению со статической сегментацией отдельных кадров, предложенный подход дает лучшее качество сегментации при том же количестве сегментов.

Качество оптического потока (табл. 3) вычислялось посредством применения формулы (4), где N соответствует числу пикселей изображения, а T_i – вектора сдвигов для каждой точки с первого кадра на второй.

Сравнительный анализ (табл. 3) показывает, что предложенные подходы к сегментации (в среднем) позволяют заметно уменьшить погрешность оптического потока. Это объясняется тем, что нахождение межкадрового движения сегмента является более ус-

тойчивой к шуму операцией, чем поиск сдвига отдельной точки.

Табл. 3. Оценка качества для исходных ($Farn$) [12] и сегментированных (DSS , DAS) оптических потоков

Метод \ Кадр	$Farn$	DSS	DAS
000.bmp	9,95	3,36	2,89
004.bmp	16,74	8,64	5,09
300.bmp	11,64	8,56	6,13
500.bmp	10,86	6,03	6,55
504.bmp	2,52	3,10	1,99
600.bmp	5,36	5,14	3,71
604.bmp	6,61	5,13	4,71
616.bmp	7,04	6,42	6,78
618.bmp	5,51	4,04	4,33
619.bmp	12,27	7,14	4,25
621.bmp	7,25	4,45	4,15
Ср. Q	8,70	5,64	4,6

Обычно использование аффинных преобразований позволяет получить лучшее качество потока (сегментации) по сравнению с поиском сдвигов для сегментов. Однако иногда ситуация становится противоположной. Это обусловлено меньшей устойчивостью алгоритма поиска аффинных преобразований для малых сегментов.

В первом столбце (рис. 1) – пара кадров, полученная посредством применения аффинных преобразований к фону и каждому спрайту. Во втором и третьем столбцах – пары кадров из реальных видео.

Нахождение сдвигов не всегда позволяет найти сегменты с аффинными межкадровыми движениями (фон на рис. 2а). Вместо этого будет найдена аппроксимация несколькими сегментами, которые не поддаются корректной склейке из-за различных сдвигов. Можно предположить, что переход к преобразованиям более общего вида, например проективным, позволит точнее описывать движение больших сегментов.

Заключение

Предложенный подход к слиянию сегментов на основе стоимости объединения не зависит от используемого типа межкадрового преобразования. Это позволяет использовать другие виды преобразований без существенных изменений алгоритма сегментации.

Предложенные методы склейки и уточнения границ могут выступать в качестве отдельных алгоритмов сегментации. Без пересчета межкадровых преобразований сегментов процедура уточнения границ сходится за конечное число шагов.

Работа метода существенно зависит от точности найденного оптического потока. Однако в процессе сегментации исходный поток может быть существенно улучшен.

Одним из недостатков предложенного метода является отсутствие инвариантности к межкадровым изменениям уровня освещенности. Это может быть исправлено предварительной гистограммной обработкой по усреднению распределений интенсивностей пары кадров.

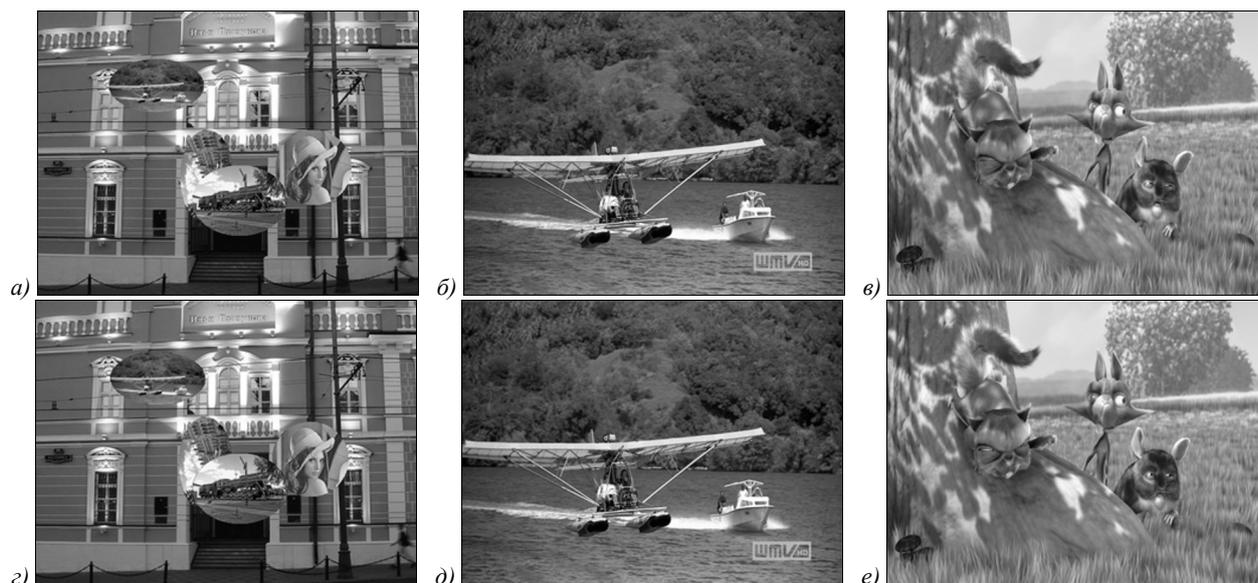


Рис. 1. Исходные пары кадров

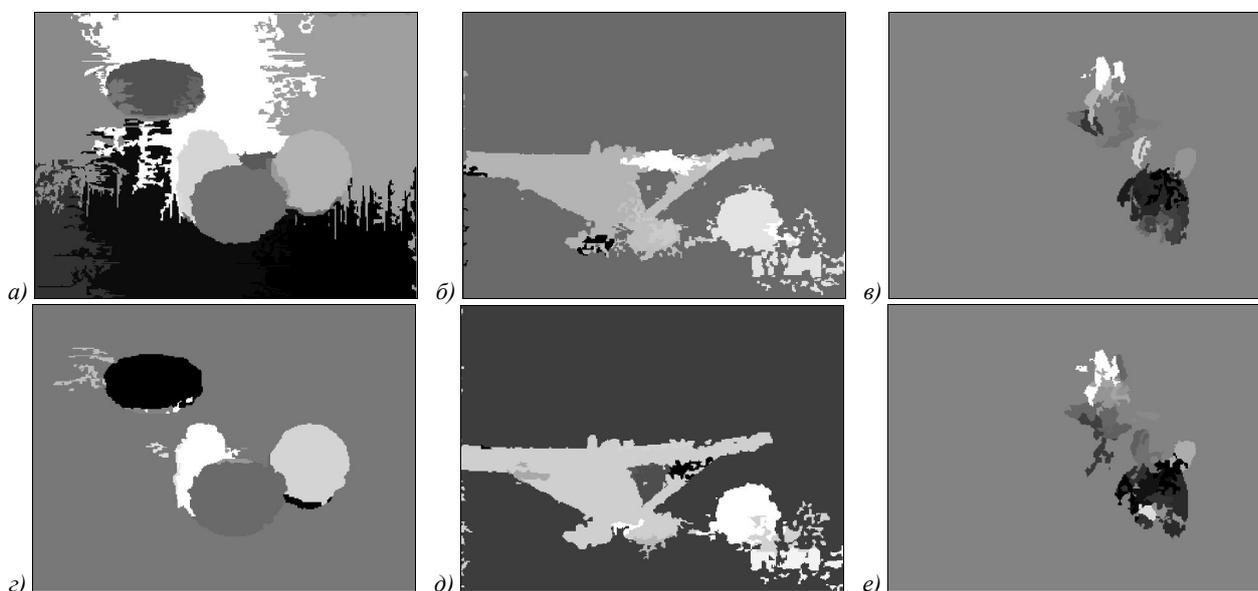


Рис. 2. Результат работы методов сегментации DSS (верхняя строка) и DAS (нижняя строка)

Подбор оптимального значения параметра α после склейки каждой пары сегментов позволит получить более высокое качество сегментации. Разработка алгоритма для решения данной проблемы требует дополнительного исследования, выходящего за рамки данной работы.

Разработанный алгоритм может быть использован для решения многих задач, в том числе: поиск и отслеживание объектов в видеопотоке, отделение подвижных объектов от фона и нахождение их границ, кодирование видео и интерполяция кадров, улучшение оптического потока и других.

Литература

1. **Ghazvinian Zanjani, F.** Improving semantic video segmentation by dynamic scene integration / F. Ghazvinian Zanjani, M. van Gerven // NCCV 2016: the Netherlands Conference on Computer Vision. – 2016. **Tsai, Y.-H.** Video segmentation via object flow / Yi-Hsuan Tsai, Ming-Hsuan Yang, Michael J. Black // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2016. – P. 3899-3908. – DOI: 10.1109/CVPR.2016.423.
2. **Kumar, M.P.** Learning layered motion segmentations of video / M.P. Kumar, P.H.S. Torr, A. Zisserman // International Journal of Computer Vision. – 2008. – Vol. 76, Issue 3. – P. 301-319. – DOI: 10.1007/s11263-007-0064-x.
3. **Sevilla-Lara, L.** Optical flow with semantic segmentation and localized layers / L. Sevilla-Lara, D. Sun, V. Jampani, M.J. Black // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2016. – P. 3889-3898. – DOI: 10.1109/CVPR.2016.422.
4. **Zappella, L.** New trends in motion segmentation / L. Zappella, X. Llado, J. Salvi. – In: Pattern Recognition / ed. by P.-Y. Yin. – Rijeka, Croatia: IntechOpen, 2009. – Chapter 3. – P. 31-46. – DOI: 10.5772/7551.
5. Optical flow evaluation results [Electronical Resource]. – URL: <http://vision.middlebury.edu/flow/eval/results/results-i1.php> (request date 22.10.2018).

6. **Lucas, B.D.** An iterative image registration technique with an application to stereo vision / B.D. Lucas, T. Kanade // Proceedings of the 1981 DARPA Image Understanding Workshop. – 1981. – P. 121-130.
7. **Baker, S.** Lucas-Kanade 20 years on: A unifying framework / S. Baker, I. Matthews // International Journal of Computer Vision. – 2004. – Vol. 56, Issue 3. – P. 221-255. – DOI: 10.1023/B:VISI.0000011205.11775.fd.
8. **Хашин, С.И.** Оценка качества сегментации изображения // Вестник Ивановского государственного университета. Серия: естественные, общественные науки. – 2010. – № 2. – С. 112-117.
9. **Szeliski, R.** Computer vision: Algorithms and applications / R. Szeliski. – London: Springer, 2011. – 824 p. – Chapter 8. – ISBN: 978-1-84882-934-3.
10. **Ваганов, С.Е.** Сравнение эффективности различных версий метода Лукаса-Канаде / С.Е. Ваганов, С.И. Хашин // Вестник Ивановского государственного университета. Серия "Естественные, общественные науки". – 2017. – Вып. 2. – С. 63-70.
11. **Farneback, G.** Two-frame motion estimation based on polynomial expansion / G. Farneback. – In: Image analysis. SCIA: Scandinavian conference on image analysis / ed. by J. Bigun, T. Gustavsson. – Berlin, Heidelberg: Springer, 2003. – P. 363-370. – DOI: 10.1007/3-540-45103-X_50.
12. **Хашин, С.И.** Аффинная версия алгоритма Лукаса–Канады // Математические методы распознавания образов: Доклады XV Всероссийской конференции. – 2011. – С. 459-462.
13. Тестовые пары bmp-файлов [Электронный ресурс]. – URL: http://math.ivanovo.ac.ru/dalgebra/Khashin/bmp_ex2/ (дата обращения 22.10.2018).
14. **Кручинин, А.Е.** Сегментация изображения путём выделения непрерывных границ / А.Е. Кручинин, С.И. Хашин // Вестник Ивановского государственного университета. – 2007. – Вып. 3. – С. 80-83.
15. **Хашин, С.И.** Динамическая сегментация последовательности кадров / С.И. Хашин // Машинное обучение и анализ данных. – 2013. – Т. 1, № 6. – С. 787-795.

Сведения об авторах

Ваганов Сергей Евгеньевич, 1992 года рождения, в 2015 окончил магистратуру в Ивановском государственном университете по направлению «Математика и компьютерные науки». В настоящий момент (2018 год) является аспирантом Ивановского государственного университета (ИвГУ) по профилю 01.06.01 «Математическая логика, алгебра и теория чисел», работает ведущим программистом факультета математики и компьютерных наук в ИвГУ. Область научных интересов: цифровая обработка изображений, распознавание образов, компьютерное зрение, машинное обучение и численные методы. E-mail: pro100-pioner@mail.ru.

ГРНТИ: 27.41.23

Поступила в редакцию 22 июня 2018 г. Окончательный вариант – 14 ноября 2018 г.

A method for dynamic segmentation of a pair of sequential video-frames

S.E. Vaganov¹

¹Ivanovo State University, Ivanovo, Russia

Abstract

An algorithm of dynamic segmentation of sequential frame pairs was proposed. A comparative analysis of segmentation quality when finding shifts and affine inter-frame transformations for the segments was conducted. In addition, we compared the performance of the proposed method with some static segmentation approaches.

Keywords: segmentation, image, video, affine transformation, optical flow.

Citation: Vaganov SE. A method for dynamic segmentation of a pair of sequential video-frames. Computer Optics 2019; 43(1): 83-89. DOI: 10.18287/2412-6179-2019-43-1-83-89.

References

- [1] Ghazvinian Zanjani F, van Gerven M. Improving semantic video segmentation by dynamic scene integration. The Netherlands Conference on Computer Vision 2016.
- [2] Tsai Y-H, Yang M-H, Black MJ. Video segmentation via object flow. IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2016: 3899-3908. DOI: 10.1109/CVPR.2016.423.
- [3] Kumar MP, Torr PHS, Zisserman A. Learning layered motion segmentations of video. International Journal of Computer Vision 2008; 76(3): 301-319. DOI: 10.1007/s11263-007-0064-x.
- [4] Sevilla-Lara L, Sun D, Jampani V, Black MJ. Optical flow with semantic segmentation and localized layers. IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2016: 3889-3898. DOI: 10.1109/CVPR.2016.422.
- [5] Zappella L, Llado X, Salvi J. New trends in motion segmentation. In Book: Yin P-Y, ed. Pattern Recognition. Ch 3. Rijeka, Croatia: IntechOpen; 2009: 31-46. DOI: 10.5772/7551.
- [6] Optical flow evaluation results. Source: <http://vision.middlebury.edu/flow/eval/results/results-i1.php>.
- [7] Lucas BD, Kanade T. An iterative image registration technique with an application to stereo vision. Proceedings of the 1981 DARPA Image Understanding Workshop 1981: 121-130.
- [8] Baker S, Matthews I. Lucas-Kanade 20 years on: A unifying framework. International Journal of Computer Vision 2004; 56(3): 221-255. DOI: 10.1023/B:VISI.0000011205.11775.fd.
- [9] Hashin SI. The estimation of quality of segmentation of the map [In Russian]. Ivanovo State University bulletin, Series "Natural, Social Sciences" 2010; 2; 112-117.
- [10] Szeliski R. Computer vision: Algorithms and applications. London: Springer; 2011. ISBN: 978-1-84882-934-3.

- [11] Vaganov SE, Hashin SI. A comparison of the effectiveness of different versions of Lucas–Kanade method [In Russian]. Ivanovo State University Bulletin, Series “Natural, Social Sciences” 2017; 2: 63-70.
- [12] Farneback G. Two-frame motion estimation based on polynomial expansion. In Book: Bigun J, Gustavsson T, eds. Image analysis. SCIA: Scandinavian conference on image analysis. Berlin, Heidelberg: Springer; 2003: 363-370. DOI: 10.1007/3-540-45103-X_50.
- [13] Khashin SI. Affine version of Lucas–Kanade algorithm [In Russian]. Proceedings of Mathematical Method of Pattern Recognition 2011: 459-462.
- [14] Test bmp-file pairs [In Russian]. Source: http://math.ivanovo.ac.ru/dalgebra/Khashin/bmp_ex2/.
- [15] Kruchinin AE, Hashin SI. The image segmentation by building continuous borders [In Russian]. Ivanovo State University Bulletin 2007; 3: 80-83.
- [16] Khashin SI. Dynamic segmentation of frames sequences [In Russian]. Machine Learning and Data Analysis 2013; 6(1): 787-795.

Author’s information

Sergey Evgenevich Vaganov (b. 1992) graduated from Ivanovo State University in 2015 (Master’s degree in Mathematics and Computer Science). Currently, he is a post-graduate student of Ivanovo State University majoring in «Mathematical Logic, Algebra and Number Theory». Also, he works as a leading programmer at the Mathematic and Computer Science department of Ivanovo State University. His main research interests are digital image processing, pattern recognition, computer vision, machine learning and numerical methods. E-mail: pro100-pioner@mail.ru.

Received June 22, 2018. The final version – November 14, 2018.
