

И. С. РУБИНА, А. Ю. ТРОПЧЕНКО

ИССЛЕДОВАНИЕ АЛГОРИТМОВ ВЫБОРА ОПОРНЫХ ПИКСЕЛОВ В ЗАДАЧАХ ВЫДЕЛЕНИЯ СЕГМЕНТОВ КАДРА ВИДЕОПОСЛЕДОВАТЕЛЬНОСТИ

Предложена схема выделения опорных пикселей, основанная на исследовании существующих алгоритмов и возможности их применения в задачах классификации сегментов кадра видеопоследовательности.

Ключевые слова: сегментный подход, блок переменного размера, опорные точки, классификация сегментов, компенсация движения.

Введение. Задача классификации сегментов кадра возникает при компенсации движения объектов в целях устранения временной избыточности видеоряда, предполагающей наличие подобия соседних кадров видеопоследовательности. Кроме того, разделение сегментов на группы в соответствии с некоторым признаком позволяет осуществить приближение к объектному подходу, что обеспечивает возможность более гибкого манипулирования содержимым видеоряда, а также решения таких прикладных задач, как выявление факта движения и определение траектории движущихся объектов.

Выбор данного подхода к компенсации движения объектов естественных видеосцен обусловлен возможностью устранения значительного количества недостатков решений, основанных на попиксельном подходе, характеризующемся высокой трудоемкостью алгоритма, и объектном подходе, отличающемся сложностью определения формы объекта прогнозирования.

Исходными данными для подобного рода задач являются прямоугольные блоки (сегменты), каждому из которых соответствует вектор движения, устанавливающий связь блоков двух соседних кадров с максимальным значением коэффициента их соответствия и определяющий параметры параллельного переноса. Разделяя объекты на классы, можно в значительной степени сократить количество векторов движения. Необходимо заметить, что согласно ряду исследований выбор критерия соответствия не влияет существенно на качество компенсации движения [1], в связи с чем следует руководствоваться требованием минимума вычислительных затрат.

Подход к устранению временной избыточности видеопоследовательности на основе блоков можно реализовать при помощи сегментов фиксированного и переменного размера. При этом выбор размера блока всегда является компромиссом, так как использование большого количества мелких блоков позволяет более подробно описывать движение в рамках кадра посредством уменьшения степени сжатия за счет увеличения количества передаваемой

информации о движении. Алгоритм компенсации движения на основе блоков переменного размера позволяет нивелировать данное ограничение за счет применения малых блоков для областей с мелкими деталями и динамичным движением и больших блоков для однородных областей типа „фон“. Как показали эксперименты, применение этого алгоритма для одной и той же видеопоследовательности обеспечивает сокращение потока векторов движения на 60 % по сравнению с подходом, основанным на использовании блоков фиксированного размера [2].

Однако даже в рамках данного подхода возникает ряд проблем. Для блоков большого размера выбор размера сегмента осуществляется по принципу мажоритарности, что способствует поглощению малых областей плавного движения. Кроме того, в связи с небольшим количеством анализируемых пикселей для блоков малого размера возможно возникновение ложных соответствий. Решить первую проблему можно, исключив использование блоков большого размера, а вторую — с помощью классификации блоков в соответствии с некоторым признаком.

В настоящей статье в качестве такого признака предлагается использовать фактор принадлежности большинства пикселей блока некоторой области, определяемой маской, построенной по представленному набору опорных точек.

Схемы алгоритмов. В ходе исследования были проанализированы существующие схемы выделения опорных пикселей на кадрах видеопоследовательности относительно их применимости к решению задач компенсации движения объекта. В настоящее время наибольшее распространение получили следующие алгоритмы:

— алгоритм ADC (Absolute Difference Criteria), в котором используется критерий абсолютного отклонения; данный показатель соответствия отличается минимальными вычислительными затратами, однако в ходе экспериментов было выявлено, что этот алгоритм выделяет неоправданно большое количество опорных точек, зависящее от перепадов освещенности сцены, а также от равномерности представленной текстуры кадра;

— алгоритм детектирования Харриса, описанный в работе [3], при высокой трудоемкости реализации не позволяет выделить стабильный набор точек;

— алгоритм SIFT (Scale-Invariant Feature Transform), описанный в работе [4], в дополнение к алгоритму ADC использует предварительную линейную свертку изображения с ядром Гаусса:

$$y(i, j) = \frac{1}{2\pi\sigma^2} \sum_{k=-m}^m \sum_{n=-m}^m e^{-\frac{k^2+n^2}{2\sigma}} y'(i+k, j+n),$$

где y' и y — значения исходной и сглаженной яркости пикселей в точках кадра с координатами (i, j) ; $(2m+1) \times (2m+1)$ — размер ядра Гаусса; σ — параметр отклонения фильтра.

Исследования проводились для разных размеров ядра Гаусса при $\sigma = 1$. При использовании ядра размером 5×5 была зафиксирована наилучшая способность фильтрации шумов, что обеспечило подходящий набор опорных точек.

По результатам проведенного анализа алгоритм SIFT был выбран в качестве основы для расчета маски классификации.

С учетом выбранной схемы был разработан следующий алгоритм построения маски классификации.

1. Выделение опорных точек по алгоритму SIFT.
2. Создание локальной окрестности — зоны интереса — для каждой точки: границы строки интереса, а также левых и правых границ изображения.
3. Создание массива строк интереса из строк длиной более 2 пкс.
4. Объединение строк в зоны интереса по принципу их стыковки со смещением в 2 пкс сверху или снизу.

5. Создание второго массива строк интереса из коротких строк в областях резких перепадов яркости, а также строк, составленных из опорных точек и точек зоны интереса.

6. Объединение коротких строк в зоны интереса по принципу их стыковки без смещения сверху или снизу.

Данный алгоритм основан на механизме „неслияния“ зон интереса через их границу, имеющую ширину в 1 пкс. Благодаря объединению коротких строк по указанной схеме происходит различение граничных строк обрамления и областей с перепадами яркости. Для уменьшения трудоемкости алгоритма предлагается его реализацию осуществлять на кадре пониженного разрешения, что, как показало исследование, практически не сказывается на качестве классификации. Стартовые кадры видеопоследовательностей (а), а также соответствующие результаты построения масок (б) представлены на рис. 1.

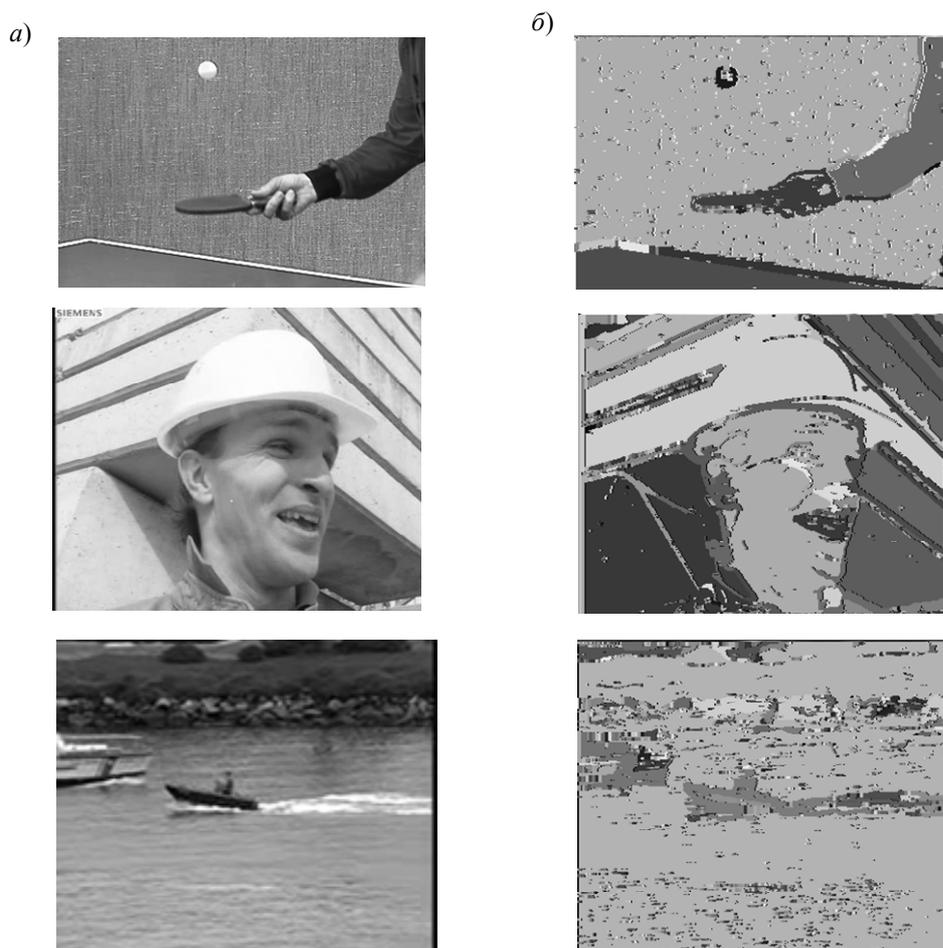


Рис. 1

Полученная в результате выполнения алгоритма маска используется в алгоритме компенсации движения на основе блоков переменного размера (VSBM — Variable Size Block Compensation) с точностью 0,25 пкс. Основные этапы такого алгоритма сводятся к следующим действиям.

1. Изображение делится на блоки равного размера.
2. Для каждого блока выполняется проверка условия: если ошибка соответствия обрабатываемого блока и наиболее подходящего блока опорного кадра выше некоторого порога, то блок делится на четыре блока меньшего размера.
3. При достижении максимально заданного числа блоков или требуемой точности совпадения кадров процесс для данного блока останавливается.

Процесс соответствия блоков зонам интереса выполняется по принципу мажоритарности в зависимости от значения маски пикселей, принадлежащих блоку. В алгоритме один вектор движения присваивается группе сегментов, принадлежащих одной зоне интереса. Результаты выполнения предложенного алгоритма представлены на рис. 2.

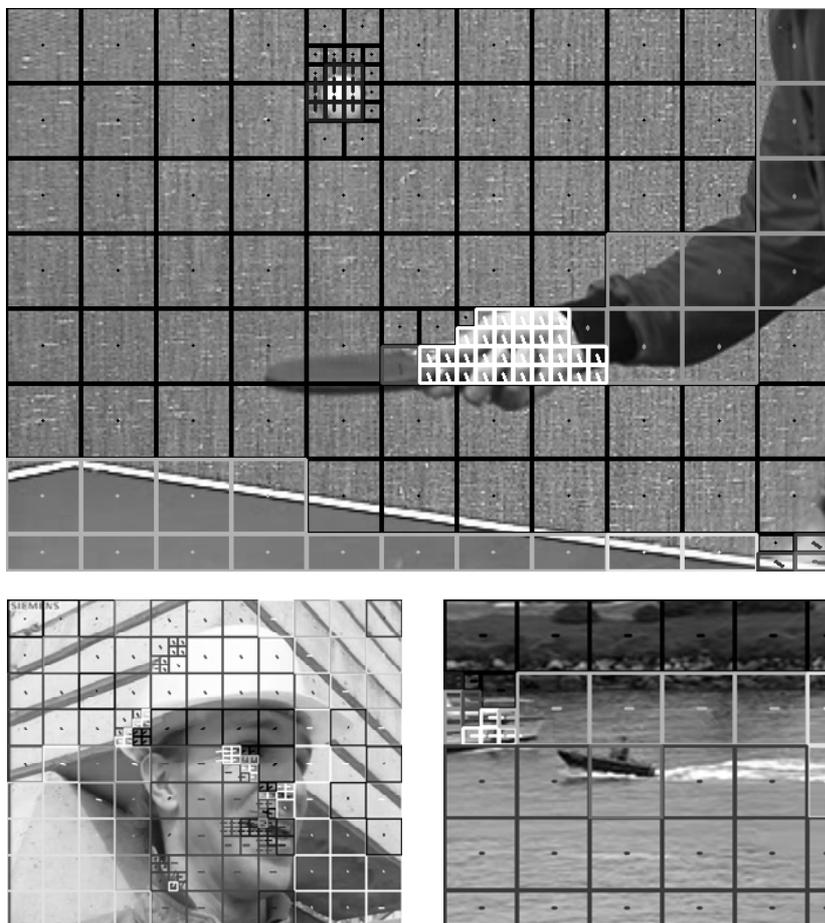


Рис. 2

В связи с трудоемкостью алгоритма полного перебора целесообразно применение метода усредненной пирамиды (MP — Median Pyramid) для отбора соотносимых блоков, который требует меньших вычислительных затрат при схожем качестве восстановленной видеопоследовательности. Алгоритм MP предполагает осуществление компенсации движения на кадре пониженного разрешения с дальнейшим уточнением векторов на кадре более высокого разрешения [5]. В рамках данного исследования предлагается использовать пирамиду Гаусса с тремя ступенями.

Экспериментальные результаты. Для тестирования разработанного алгоритма были выбраны стандартные тестовые последовательности группы MPEG — „Теннис“, „Бригадир“ и „Береговая охрана“.

В ходе экспериментов для алгоритма VSBM и модифицированных алгоритмов VSBM на основе выделения опорных пикселей (VSBMP) и на основе схемы отбора блоков (VSBMPM) были получены следующие зависимости:

— пиковое соотношение сигнал/шум для всего перечня размеров блока, вычисляемое в соответствии с формулой

$$\text{PSNR} = 10 \lg \frac{(2^n - 1)^2}{\text{MSE}},$$

где n — разрядность цветовой схемы, для цветового пространства YUV равная 8; MSE — среднее квадратическое отклонение яркости пикселей исходного изображения от яркости изображения, восстановленного после сжатия, определяемое как

$$\text{MSE} = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \|I(i, j) - K(i, j)\|^2,$$

где M, N — размеры кадра; $I(i, j)$ и $K(i, j)$ — яркость соответствующих пикселов компенсируемого и опорного кадров;

— RD (Ration Distortion) — характеристика, определяющая зависимость искажения сигнала (PSNR) от коэффициента сжатия ($R(D)$);

— вычислительная сложность алгоритма (Q), измеряемая как среднее количество базовых операций (б.о.) (сложения/вычитания и сравнения) на блок для всего перечня его размеров.

Для тестовой последовательности „Теннис“ на рис. 3, а представлена зависимость величины PSNR от площади S блока выделяемого сегмента, которая увеличивается с шагом d по каждой из сторон; на рис. 3, б приведена зависимость вычислительной сложности алгоритма Q от величин S и d , а на рис. 3, в — зависимость $R(D)$.

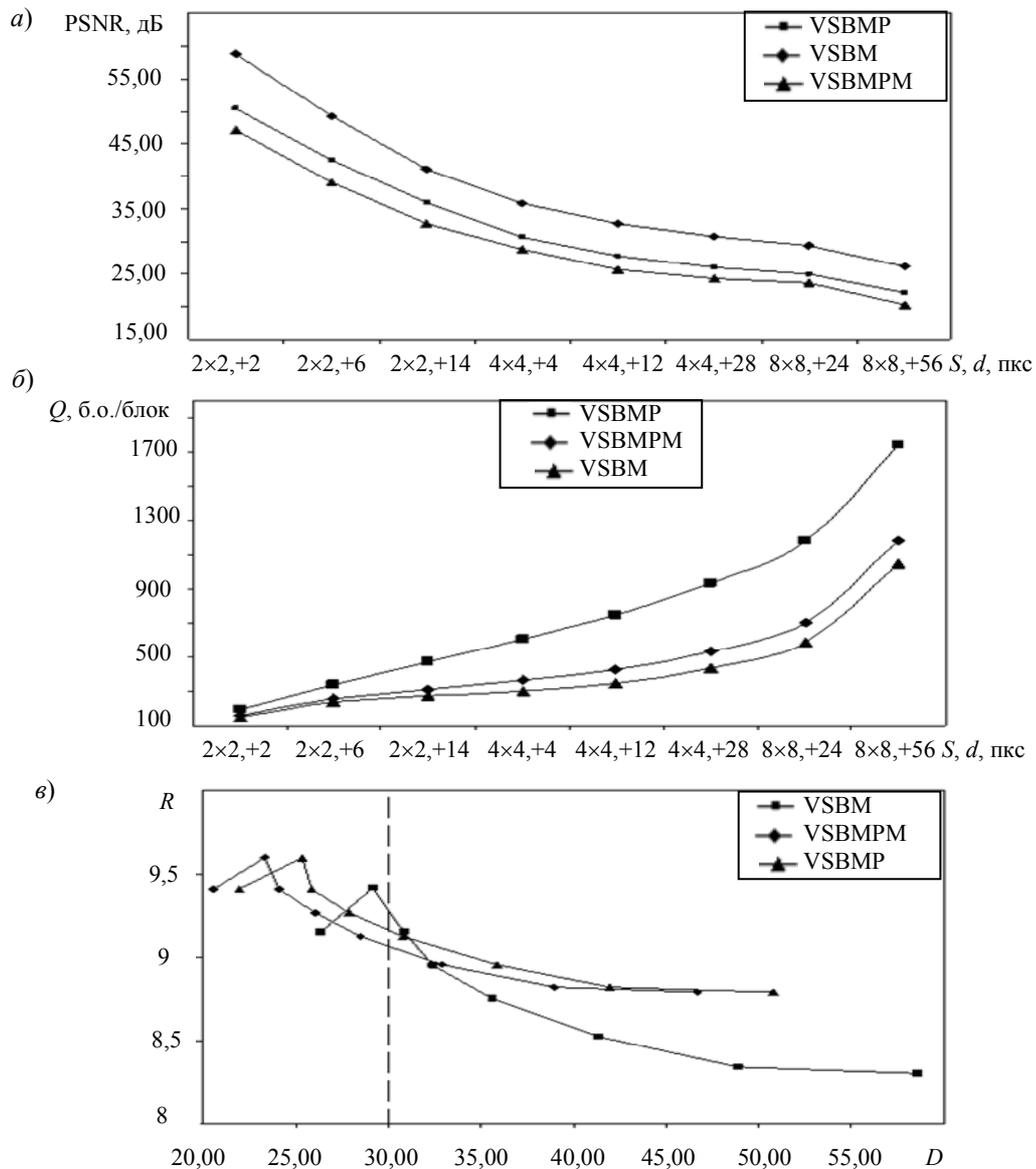


Рис. 3

Анализ результатов, представленных на рис. 3, а также в таблице, показал, что предложенный алгоритм способствует улучшению показателей сжатия при несущественном понижении качества воспроизведенной видеопоследовательности (если PSNR > 30дБ, то качество

считается хорошим). При использовании алгоритма VSBMPM было достигнуто снижение вычислительной трудоемкости при схожих показателях коэффициента сжатия и незначительном ухудшении качества. Необходимо также заметить, что при максимальном значении размера блока, равном 64×64 , проявляется эффект мажоритарности.

Размер блока, пкс		Исследуемый алгоритм								
		VSBM			VSBMP			VSMPPM		
min	max	PSNR, дБ	$R(D)$	Q , б.о.	PSNR, дБ	$R(D)$	Q , б.о.	PSNR, дБ	$R(D)$	Q , б.о.
2×2	4×4	58,3	8,30	153	50,5	8,79	206	46,6	8,79	163
	8×8	48,7	8,34	240	42,1	8,82	342	39,4	8,82	260
	16×16	41,9	8,52	271	35,7	8,96	473	32,9	8,96	311
4×4	8×8	35,8	8,76	307	30,7	9,13	605	28,5	9,13	367
	16×16	32,9	8,95	359	28	9,27	758	26	9,27	439
	32×32	30,7	9,15	440	26	9,41	937	24,3	9,41	540
8×8	32×32	29,9	9,42	585	25,6	9,60	1187	23,2	9,60	705
	64×64	25,8	9,15	1048	22	9,41	1751	20,3	9,41	1188

Заключение. В ходе исследования разработана схема выделения локальных областей кадра (зон интереса), основанная на методе SIFT выбора опорных точек, отличающаяся уменьшением вычислительных затрат и допустимым уровнем точности с возможностью использования в задачах классификации сегментов кадра.

Применение предложенной схемы обеспечивает более высокий коэффициент сжатия видеопоследовательности, чем в схеме VSBM, но при более значимых вычислительных затратах. Этот показатель можно улучшить, как было показано, за счет применения алгоритма усредненной пирамиды отбора блоков для сравнения.

СПИСОК ЛИТЕРАТУРЫ

1. *Srinivasan R., Rao K. R.* Predictive coding based on efficient motion estimation // Proc. of ICC. 1984. P. 521—526.
2. *Ribas-Corbera J., Neuhoff D. L.* On the optimal block size for block-based, motion compensated video coders // SPIE Proc. of Visual Communications and Image Processing. 1997. Vol. 3024. P. 1132—1143.
3. *Harris C., Stephens M.* Combined corner and edge detector // Proc. of the 4th Alvey Vision Conference. 1988. P. 147—151.
4. *Lowe D. G.* Object recognition from local scale-invariant features // Proc. of the Intern. Conf. on Computer Vision. 1999. Vol. 2. P. 1150—1157.
5. *Рубина И. С.* Анализ методов построения траектории движущихся объектов на основе сегментации видеоданных // Науч.-техн. вестн. СПбГУ ИТМО. 2011. Вып. 72.

Сведения об авторах

Ирина Семеновна Рубина

— аспирант; Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кафедра вычислительной техники; E-mail: rubren@mail.ru

Александр Ювенальевич Тропченко

— д-р техн. наук, профессор; Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кафедра вычислительной техники; E-mail: tau@dl.ifmo.ru

Рекомендована кафедрой
вычислительной техники

Поступила в редакцию
03.03.11 г.