

ЦИФРОВАЯ СТАБИЛИЗАЦИЯ ИЗОБРАЖЕНИЙ В УСЛОВИЯХ ЗАПЛАНИРОВАННОГО ДВИЖЕНИЯ

© 2012 г. С. П. Званцев*, канд. техн. наук; П. И. Иванов**, студент;
Е. Ю. Мерзлютин*, канд. техн. наук

* ЗАО НИИВЦ “Карат”, Санкт-Петербург

** Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, Санкт-Петербург

E-mail: pavel.ivanov.lp@gmail.com

Рассмотрены аспекты решения задачи стабилизации изображений при съемке подвижной камерой для улучшения восприятия изображений сцены оператором в условиях, когда размеры сцены превышают размеры области визуализации. Предложены методы повышения качества выбора пар точек для оценки параметров реального и запланированного движения изображения сцены. Обсуждается подход к реализации алгоритма в реальном времени.

Ключевые слова: цифровая стабилизация изображений, подвижное основание, оценка параметров движения.

Коды OCIS: 100.2960

Поступила в редакцию 29.05.2012

Введение

Стабилизация изображения – актуальная задача при обработке видеопоследовательностей, получаемых от видеокамер или оптико-электронных систем (ОЭС), установленных на подвижном основании (носителе). Съемка может вестись как с транспортного средства, так и “с руки” оператором. В обоих случаях присутствуют случайные движения поля зрения камеры относительно объекта съемки, вызванные угловыми и линейными смещениями носителя. Стабилизация изображений сцены тесно связана с визуализацией изображения на мониторе оператора в процессе наблюдения. Область изображения, предназначенную для отображения оператору, будем называть областью визуализации. Область визуализации содержит информацию, полученную из текущего кадра, в том числе после стабилизации изображения относительно заданного направления поля зрения. При этом текущий кадр в его пространственной ориентации можно отождествлять с данными датчика изображения. Под стабилизацией изображений понимается комплекс мер, обеспечивающих комфортное восприятие оператором сцены в области визуализации на мониторе в условиях случайного

движения поля зрения ОЭС. В целях улучшения восприятия информации при регулярном движении поля зрения относительно сцены стабилизация может осуществляться путем компенсации случайной составляющей этого движения с сохранением систематической составляющей. Отсутствие стабилизации, помимо ухудшения зрительного восприятия сцены, усложняет решение задач обработки изображений, таких как обнаружение изображений подвижных объектов на фонах и их межкадровой идентификации при сопровождении или поиске корреляционного подобия.

Известен ряд подходов к стабилизации изображений, в том числе применение средств механической, оптической или электронной стабилизации. В настоящее время значительный интерес вызывают методы цифровой стабилизации изображений [1–7].

В гиросtabilизированных ОЭС, как правило, выполняется двухосная стабилизация поля зрения, при которой изображение движется относительно строк матрицы камеры в соответствии с движением центра масс носителя до появления ненулевой проекции угловой скорости ориентации носителя в направлении оптической оси ОЭС. Ее наличие приводит к вращению изображения на мониторе оператора. Для

компенсации вращения может использоваться дополнительная стабилизация по третьей оси, хотя, зачастую, от нее отказываются ввиду сложности реализации и появления у оператора неадекватного представления о его собственной ориентации в пространстве при длительном наблюдении. Так же при гироскопической стабилизации (как и при электронной) отсутствует стабилизация масштаба изображения.

В отличие от гиросtabilизации, цифровая стабилизация связана с наблюдаемой сценой, что имеет как достоинства, так и недостатки. К достоинствам можно отнести возможность переключения режимов стабилизации (например, компенсации только вращения или масштаба изображения, фиксации положения сцены, фиксации с периодической сменой опорного положения сцены и т. п.). Использование цифровых методов в дополнение к двухосной гироскопической стабилизации позволяет обеспечить ситуационное включение стабилизации по третьей оси на время изучения сцены.

Цифровые методы позволяют уменьшить габариты и стоимость оптико-электронной системы за счет отказа от использования гиросtabilизированных платформ или снижения требований к ним, что является актуальным для беспилотных летательных аппаратов и малогабаритных мобильных роботов. Цифровая стабилизация незаменима, когда требуется осуществить стабилизацию ранее снятого видео, т. е. выполнить его пост-обработку.

К недостаткам цифровых методов можно отнести: наличие методических ошибок, вызванных неадекватностью применяемых моделей преобразования изображений, приводящих к накоплению ошибок в оценках параметров преобразований; сложность реализации алгоритмов в реальном времени; технические проблемы, связанные с передачей видеопотока с малым коэффициентом сжатия к компьютеру обработки; ограниченный выбор аппаратных средств, позволяющих выполнять параллельные вычисления и удовлетворяющих требованиям бортового применения.

При наблюдении оператором с подвижно-го основания можно выделить несколько этапов: обзор пространства, наблюдение объекта (в том числе при удержании области интереса средствами сопровождения), наведение на выбранный объект. Стабилизация полезна для первых двух этапов, поскольку в них происходит визуальный поиск области интереса и анализ сцены, требующие движения изо-

бражения, удобного для принятия решения. На этапе ручного наведения на объект (например, с помощью джойстика) стабилизация должна быть отключена для обеспечения визуальной обратной связи со сценой по изображению на мониторе. Если цель однократно помечена оператором на мониторе, а наведение поля зрения осуществляется автоматически, то стабилизация в виде плавного наведения оптической оси на цель возможна и на этом этапе.

Преобразования координат при стабилизации

Суть рассматриваемых методов стабилизации состоит в получении стабилизированного изображения в области визуализации, отображаемой на мониторе оператора или в поле обработки алгоритмов обнаружения и сопровождения подвижного объекта. В момент включения стабилизации фиксируется область визуализации, которая совпадает с текущим кадром. При получении следующего кадра путем оценки параметров межкадрового преобразования изображения W_1 появляется возможность попиксельного отображения изображения текущего кадра в область визуализации. Процесс расчета межкадровых преобразований W_k и суммарного преобразования

$$\bar{W}_i = \prod_{k=1}^i W_k \quad (1)$$

повторяется для последующих i -х кадров. Этот процесс описывает стабилизацию изображения в области визуализации. Дальнейшая работа по стабилизации основана на следующих рассуждениях.

Получаемые текущие изображения являются двумерными проекциями трехмерных сцен, наблюдаемых при линейных и угловых движениях камеры. Как будет показано ниже, в межкадровых преобразованиях используется набор параметров, включающий сдвиги по осям кадра, поворот, масштаб, искажения по осям, что характерно для преобразований на плоскости. Таким образом, можно рассматривать двумерную прямоугольную систему координат, связанную с начальной областью визуализации (СКОВ₀), относительно которой получают изображения смещенных, повернутых и прочими способами преобразованных текущих кадров.

При удалении текущего поля зрения от первоначального положения целесообразно

смещать область визуализации относительно $СКОВ_0$ на основе оценок параметров преобразований \bar{W}_i таким образом, чтобы движение стабилизированного изображения в текущем j -м положении области визуализации $СКОВ_j$ соответствовало плавному движению камеры относительно сцены при условии нахождения значительной доли кадра в области визуализации. Данный процесс описывает “запланированное движение”, прогнозирующее намерение оператора при осмотре сцены или предоставляющее оператору возможность ее комфортного рассмотрения при анализе.

Таким образом, преобразование изображения i -го кадра при стабилизации с учетом запланированного движения области визуализации имеет вид

$$\Omega_i = \bar{W}_i^{-1} \bar{V}_j, \quad (2)$$

$$\bar{V}_j = \prod_{n=1}^j V_n, \quad (3)$$

где \bar{V}_j преобразование координат области визуализации после j -го перемещения относительно $СКОВ_0$. Преобразования V_n могут включать смещения по осям X и Y , вращение и изменение масштаба изображения.

Отметим случайный характер преобразований W_i , полученных на основании оценок параметров, и детерминированный характер V_j , заданных на основе изменения параметров движения поля зрения относительно области визуализации. Ошибка оценки преобразований W_i отражается в ошибке стабилизации и определяется как отклонение центра кадра стабилизированного изображения от текущего положения области визуализации. Эта ошибка увеличивается с ростом числа кадров. Ошибка выбора детерминированной составляющей влияет на величину пересечения текущего кадра с областью визуализации.

В зависимости от поставленной перед оператором задачи возможны различные стратегии компенсации движения камеры, отличающиеся поведением области визуализации:

- фиксация области визуализации в выбранном направлении для статичных сцен,
- сглаженное движение области визуализации по оценкам параметров смещения и поворота, например, при панорамном обзоре, рис. 1а,
- периодическая фиксация области визуализации с последующим быстрым переходом к следующему неподвижному положению, рис. 1б, как вариант компромисса между двумя

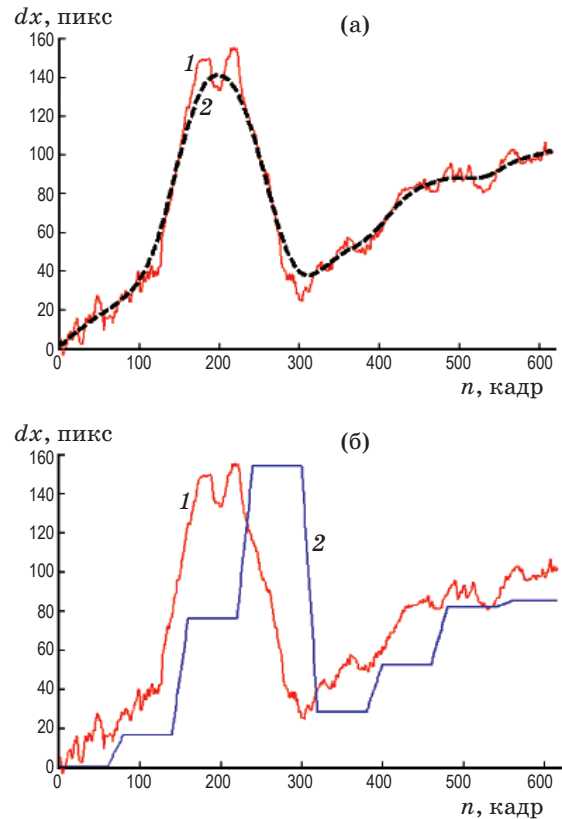


Рис. 1. Стратегии компенсации движения камеры, отличающиеся поведением области визуализации. На рисунке а) представлена модель сглаженного движения области визуализации, на рисунке б) периодическая смена фиксированного положения области визуализации. Кривая 1 – оценка параметра движения сцены (смещения по оси X), кривая 2 – принятое изменение параметра движения области визуализации.

предыдущими стратегиями для сцен с трудно предсказуемым движением или как способ изучения оператором наблюдаемой сцены в текущем положении области визуализации.

Оценка параметров преобразования

Оценка параметров межкадрового преобразования изображения строится с учетом выбранного типа преобразования. В данной работе использовано аффинное преобразование

$$X' = WX, \quad W = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}, \quad (4)$$

где X и X' – координаты соответствующих точек в n -м и $(n + 1)$ -м кадрах, W – матрица аффинного преобразования. При этом t_x и t_y от-

вечают за смещение изображения вдоль осей x и y системы координат области визуализации, а элементы a_{11} , a_{12} , a_{21} , a_{22} содержат информацию о повороте, наклоне и масштабировании. Можно записать отдельные составляющие преобразования, выделяя параметры смещения

$$\mathbf{W} = \mathbf{T}\mathbf{A}, \quad \mathbf{T} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix}, \quad (5)$$

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

где \mathbf{T} – преобразование смещения, \mathbf{A} – преобразование вращения и масштаба.

Чтобы произвести оценку матрицы аффинного преобразования \mathbf{W} , применялся метод наименьших квадратов (МНК), в котором использовались пары соответствующих особых точек в соседних кадрах. Здесь искомые элементы матрицы аффинного преобразования представлены в виде 6-мерного вектора искомых параметров \mathbf{c}

$$\mathbf{q} = \mathbf{S}\mathbf{c},$$

$$\mathbf{q} = \begin{bmatrix} x'_1 \\ \vdots \\ x'_n \\ y'_1 \\ \vdots \\ y'_n \end{bmatrix}, \quad \mathbf{S} = \begin{bmatrix} x_1 & y_1 & 0 & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n & y_n & 0 & 0 & 1 & 0 \\ 0 & 0 & x_1 & y_1 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & x_n & y_n & 0 & 1 \end{bmatrix}, \quad (6)$$

$$\mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \\ t_x \\ t_y \end{bmatrix},$$

$$\mathbf{c} = (\mathbf{S}^T \times \mathbf{S})^{-1} \times (\mathbf{S}^T \times \mathbf{q}),$$

где $x_1, \dots, x_n, y_1, \dots, y_n$ и $x'_1, \dots, x'_n, y'_1, \dots, y'_n$ – координаты соответствующих точек в n -м и $(n + 1)$ -м кадрах.

Для последующей оценки динамики запланированного движения необходимо оценить параметры преобразования, такие как смещения, поворот, масштаб, наклон. В матрице межкадрового аффинного преобразования \mathbf{W} смещения выделяются с помощью выражения (5). Для выделения параметров вращения, мас-

штаба и наклона в матрице \mathbf{A} , элементы которой содержат сочетания этих параметров, необходимо выполнить ее декомпозицию. Для этого может быть использована декомпозиция методом мультипликативного разложения. Одним из его вариантов является QR -разложение [11], дающее представление матрицы \mathbf{A} (5) в виде

$$\mathbf{A} = \mathbf{Q}\mathbf{R}, \quad (7)$$

где \mathbf{Q} – ортогональная матрица, \mathbf{R} – верхняя треугольная матрица.

В результате QR -разложения матрица \mathbf{A} может быть записана в виде

$$\mathbf{A} = \begin{bmatrix} \cos(\varphi) & -\sin(\varphi) & 0 \\ \sin(\varphi) & \cos(\varphi) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} s_x & s_y \operatorname{tg}(\theta) & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (8)$$

где φ и θ – углы поворота и наклона, соответственно, а s_x и s_y – увеличение по осям x и y системы координат области визуализации.

Следует отметить, что в аффинном преобразовании подразумевается вращение относительно начала координат изображения. Однако при вращении камеры оператором относительно оптической оси изображение поворачивается относительно центра поля зрения. При этом в аффинном преобразовании выполняется следующая последовательность действий: изображение поворачивается относительно начала координат, затем смещается таким образом, чтобы центры поля зрения исходного и повернутого изображения совпали. Это смещение неявно присутствует в оценке межкадрового смещения t_x и t_y в формуле (4). Аналогичные эффекты проявляются при масштабировании и наклонах. Для его исключения при оценке параметров межкадрового преобразования необходимо перейти к системе координат, связанной с центром поля зрения. Таким образом, оценка линейного движения камеры осуществляется по вычислению смещений центра поля зрения.

На основе оценки параметров движения камеры в серии кадров и принятой гипотезы о характере их изменения производится оценка случайной и систематической составляющих изменения этих параметров. Здесь под гипотезой движения подразумевается тип уравнения, описывающего изменения параметров: постоянное значение, линейное или квадратичное. Примером гипотезы изменения параметров могут быть постоянные значения масшта-

ба, углов поворота и наклона при линейном изменении смещений центра поля зрения.

Выбор пар точек для оценивания

Точность результата вычисления параметров преобразования в МНК зависит от качества сопоставления ключевых точек. Для получения ключевых точек и их сопоставления в работе использован метод SURF [8]. Возможен выбор точек и другими методами, например, алгоритмом блочного сопоставления 3DRS [9] или методом сопоставления ключевых точек SIFT [10], однако SURF обеспечивает потенциально более высокую точность находже-

ния пар точек при значительных межкадровых изменениях, чем блочные методы и работает быстрее метода SIFT.

Во избежание грубых ошибок необходимо отсеять максимальное количество недостоверных или малонадежных пар точек. Для этого в работе использовано три этапа фильтрации результатов сопоставления.

Первый этап представляет собой модификацию процедуры сопоставления ключевых точек метода SURF. В SURF поиск наилучшего кандидата обычно выполняется методом “ближайшего соседа”, для чего используется евклидово расстояние между дескрипторами точек

$$dist = \sqrt{\sum_{i=1}^4 \sum_{j=1}^4 \left((\sum dx'_{ij} - \sum dx_{ij})^2 + (\sum dy'_{ij} - \sum dy_{ij})^2 + (\sum |dx'_{ij}| - \sum |dx_{ij}|)^2 + (\sum |dy'_{ij}| - \sum |dy_{ij}|)^2 \right)}. \quad (9)$$

В случае, когда на “ближайшего соседа” точки претендуют одновременно несколько точек соседнего кадра, велика вероятность ошибочного сопоставления, поэтому подобные точки в выборе пары не используют.

Для повышения надежности выбора пар использовано усиленное правило: к парам с недостаточной надежностью сопоставления предлагается относить и случаи, когда вектор дескриптора, следующего за ближайшим соседом кандидата, удален от дескриптора исследуемой точки менее чем в t_{ratio} раз (значения порядка 0,5) по отношению к найденному “ближайшему соседу”.

Второй этап фильтрации недостоверных пар точек осуществляется методом RANSAC, который часто используется в алгоритмах сопоставления. Суть этого метода состоит в следующем: берутся сопоставленные ключевые точки исследуемой пары кадров, например n и $n + 1$, в них делается k случайных выборок из 3-х или более пар соответствующих точек. На основе каждой отдельной выборки по МНК строится своя модель преобразования из n -го в $(n + 1)$ -й кадр, называемая гипотезой. Затем все сопоставленные в исследуемой паре кадров точки проверяются на соответствие данной гипотезе, т. е. проверяется, будет ли после преобразования точка на n -м кадре отстоять от соответствующей ей точки на $(n + 1)$ -м кадре на расстояние, не превышающее t_{RANSAC} . Среди гипотез выбирается лучшая – соответствующая

наибольшему количеству сопоставлений, прошедших порог. Пары точек, не удовлетворяющие условию соответствия лучшей гипотезе, отбраковываются. Метод RANSAC успешно фильтрует грубые ошибки сопоставления, превышающие t_{RANSAC} , однако при этом ложные пары точек, удовлетворяющие этому порогу, не будут отбракованы.

Проведенные исследования показали, что целесообразно выполнить третий этап – дополнительную фильтрацию результатов сопоставления, методом, описанным в [12], использующим вычисление гистограмм локального увеличения с выбором максимально правдоподобного значения.

Пусть

$$S_{n,m} = \frac{\sqrt{(x'_n - x'_m)^2 + (y'_n - y'_m)^2}}{\sqrt{(x_n - x_m)^2 + (y_n - y_m)^2}}, \quad (10)$$

$$S_n = \arg \left\{ \max_{m \in [1; M]} \left(\text{hist} \left(S_{n,m} \right) \right) \right\},$$

$$S = \arg \left\{ \max_{n \in [1; N]} \left(\text{hist} \left(S_n \right) \right) \right\},$$

где $\text{hist}(\dots)$, $n \in [1; N]$ – гистограмма распределения масштабов, масштаб S соответствует координате максимума гистограммы. Тогда, если $S_{n,m} \in [S - t_{scale}, S + t_{scale}]$, то сопоставленные пары считаются надежными.

В отличие от RANSAC, в описанном методе нет нижнего предела порога чувствитель-

ности. Тем не менее, поскольку используются гистограммы, его точность зависит от размаха значений локального увеличения, поэтому использование метода целесообразно следом за методом RANSAC в качестве уточняющего. На практике этот этап фильтрации позволяет удалить из рассмотрения точки, принадлежащие движущимся объектам и объектам переднего плана сцены.

Реализуемость цифровой стабилизации в реальном времени

В экспериментальной части данной работы были промоделированы алгоритмы стабилизации с сохранением запланированной составляющей движения камеры, сопровождения точек по косвенным признакам и стабилизация с периодической сменой опорного положения области визуализации, также были произведены замеры времени, затрачиваемого различ-

ными этапами обработки. Эксперимент проводился на реальных видеопоследовательностях. На рис. 2 для сравнения представлены результаты работы алгоритма стабилизации и исходные кадры одной из видеопоследовательностей, использованных в эксперименте. Можно видеть, что в стабилизированном видео точка сцены (белый маркер), соответствующая центру области визуализации (черный маркер), как и остальные точки сцены, остается неподвижной вне зависимости от происходившего в исходной видеопоследовательности движения.

Для оценки вычислительных затрат на реализацию стабилизации между блоками программы были установлены счетчики времени. Усредненные результаты замеров представлены в виде диаграммы на рис. 3, из которой следует, что 90% времени обработки – это межкадровое сопоставление, т. е. работа метода SURF. Существует множество примеров успеш-

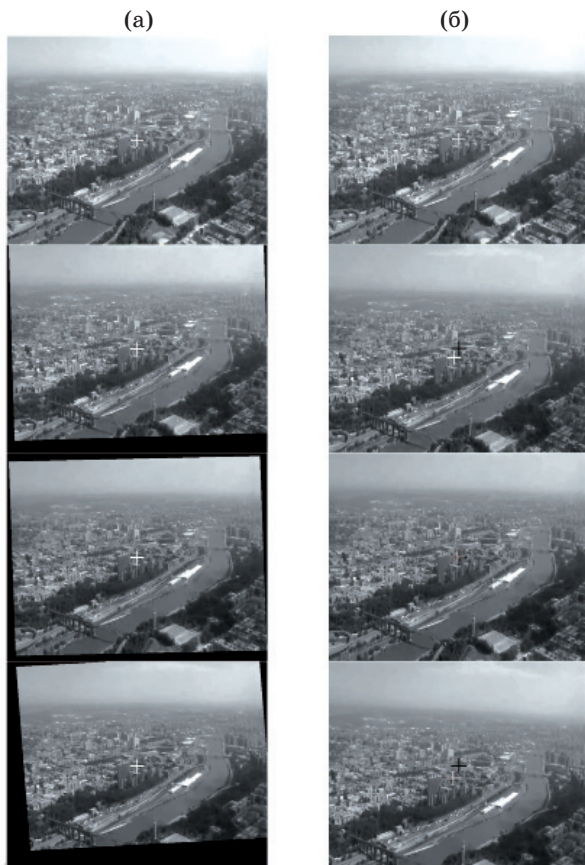


Рис. 2. Результаты обработки. В левом столбце (а) представлены: 1-й, 16-й, 28-й, 51-й кадры стабилизированной видеопоследовательности, сверху вниз, на рисунке в правом столбце (б) – кадры исходного видео.

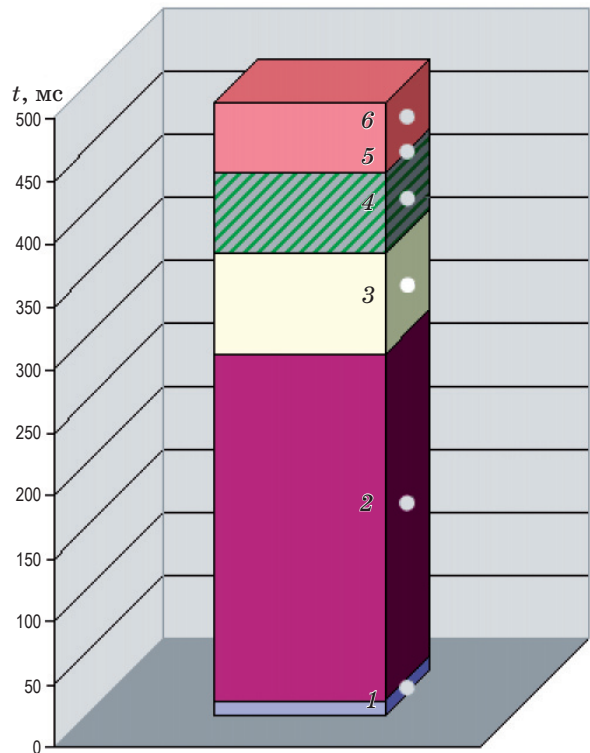


Рис. 3. Диаграмма затрат времени на реализацию стабилизации (расчет). 1 – восстановление кадра, 2 – вычисление преобразования, 3 – сопоставление точек интереса, 4 – расчет дескрипторов, 5 – выделение точек интереса, 6 – расчет интегрального изображения.

Зависимость времени обработки от размеров кадра

Размер кадра, пикс	720×576	640×480	512×384	352×288	220×176	160×112
Время обработки, мс	476	349	246	124	53	28

ной оптимизации этого метода с использованием специальных процессоров, реализующих параллельные вычисления. В частности, авторы статьи [3] достигли скорости выполнения алгоритма SURF 30 кадр/с с использованием графических процессоров NVIDIA.

Из диаграммы (рис. 3) следует, что значительная часть времени обработки кадра занята процедурами, имеющими пропорциональную зависимость времени их выполнения от площади обрабатываемого изображения. Была проведена оценка зависимости времени обработки от размера обрабатываемого фрагмента, результаты которой приведены в таблице.

Основываясь на этой зависимости, можно, уменьшая разрешение в 2 раза (пирамида изображения), увеличивать скорость обработки в 4 раза и использовать грубую оценку, инвариантную к масштабированию и поворотам, полученную методом SURF, для последующего уточнения, например, корреляционными методами.

Выводы

1. Рассмотрен подход к стабилизации видео-последовательности при съемке с подвижного основания, обеспечивающий выделение запланированной составляющей движения для повышения качества восприятия сцены оператором.

2. Предложен метод повышения надежности межкадрового сопоставления пар точек путем использования нескольких этапов проверки, в том числе за счет усиленной фильтрации неоднозначных ситуаций.

3. Оценены затраты времени на реализацию алгоритма. Показано, что 90% времени составляют линейные операции, допускающие распараллеливание вычислений. Предложено повысить скорость расчета за счет использования пирамиды изображений.

4. Проведены эксперименты, подтверждающие работоспособность и эффективность рассмотренного метода на реальных видеопоследовательностях.

* * * * *

ЛИТЕРАТУРА

1. Yang J, Schonfeld D. Robust Video Stabilization based on particle filter tracking of projected camera motion // IEEE Transection. Circuits and Systems for Video Technology. 2009. V. 19. № 7. P. 945–954.
2. Lin CT, Hong CT. Real-Time Digital Image Stabilization System Using Modified Proportional Integrated Controller // IEEE Transactions. Circuits and Systems for Video Technology. 2009. V. 19. № 3. P. 427–431.
3. Terriberry, French L.M., Helmsen J. GPU Accelerating Speeded-Up Robust Features // Proceedings of 3DPVT. 2008.
4. Pan Z., Ngo C.-W. Selective Object Stabilization for Home Video Consumers. // IEEE Transactions on Consumer Electronics. 2005. V. 51. № 4. P. 248–253.
5. Litvin A., Konrad J. Probabilistic video stabilization using Kalman filtering and mosaicking // Proc. of SPIE Electronic Imaging. 2003. V. 5022. P. 663–674.
6. Duric Z., Rosenfeld A. Shooting a smooth video with a shaky camera // Machine Vision and Applications. 2003. V. 13. P. 303–313.
7. Ertürk S. Real-time digital image stabilization using Kalman filters // Real-Time Imaging. 2002. V. 8. P. 317–328.
8. BayH., Tuytelaars T., Van Gool L. SURF: Speeded Up Robust Features // In ECCV (1). 2006. P. 404–417.

9. *De Haan G., Biezen P.W.A., Huijgen H., Ojo O.A.* True Motion Estimation with 3-D Recursive Search Block Matching // IEEE Transection o, Circuits and Systems for Video Technology. 1993. V. 3. № 6. P. 248.
 10. *Lowe D.G.* Distinctive Image Features from Scale-Invariant Keypoints // International Journal of Computer Vision. 2004.
 11. *Канатников А.Н., Крищенко А.П.* Математика в техническом университете. Выпуск IV: Линейная алгебра. М.: Изд. МГТУ им. Баумана, 2002. С. 296.
 12. *Солдатов С.А., Стрельников К.Н., Ватолин Д.С.* Быстрое и надежное определение глобального движения в видеопоследовательностях // Труды конференции “Graphicon-2006”. 2006. С. 430–437.
-