

Новый метод подавления эффекта дрожания кадра в видео

К.Н. Стрельников, Д.С. Ватолин, С.А. Солдатов

МГУ им. Ломоносова, факультет Вычислительной Математики и Кибернетики,

Лаборатория Компьютерной Графики

{strelnikov_kn, dmitriy, stass}@graphics.cs.msu.ru

Аннотация

В данной статье авторами рассмотрены наиболее известные алгоритмы подавления дрожания в видео и проанализированы их достоинства и недостатки. Также предложены новые алгоритмы подавления дрожания в видео с заполнением неизвестных области после сдвига кадра и определения наличия дрожания в видео.

1.1.1 Введение

Человеческий глаз очень чувствителен к дрожанию кадра в видео, поскольку реагирует на движение, а хаотичное движение является сильным раздражителем для глаза. Поэтому просмотр фильмов с дрожанием кадра вызывает дискомфорт восприятия. Дрожание кадра обычно появляется при съемках без использования штатива или специальных платформ, которые применяются при киносъемках. Это, как правило, архивное видео, материал журналистов или домашнее видео.

Целью алгоритмов подавления эффекта дрожания является удаление нежелательного дрожания при сохранении равномерного движения камеры.

2 Обзор существующих алгоритмов

Можно выделить четыре разных подхода к подавлению дрожания кадра:

- 1) на основе алгоритма определения движения (Motion Estimation);
- 2) на основе аппарата особых точек (Feature Points);
- 3) на основе визуализации по изображениям (Image-Based Rendering);
- 4) на основе анализа кадра целиком.

Далее будут рассмотрены несколько алгоритмов, использующих эти подходы.

2.1 Подавление дрожания в видео с использованием Motion Estimation

Аппарат Motion Estimation используется практически во всех областях обработки видео. Задача – определить направление и скорость движения небольших частей кадра. Для этого кадр разбивается прямоугольной сеткой на блоки заданного размера, и для каждого блока минимизируется функция (1).

$$error = \sum_{(\Delta x, \Delta y) \in S} (I_T(x, y) - I_{T'}(x + \Delta x, y + \Delta y))^2 \quad (1)$$

Таким образом, каждому блоку ставится в соответствие вектор движения $(\Delta x, \Delta y)$ из некоторой области поиска S . А для пары кадров определяется векторное поле, характеризующее смещение объектов сцены за время $T - T'$.

2.1.1 Простейший подход

Простейший подход [1] заключается в определении смещения кадра, как медианного вектора из найденного поля векторов. Таким образом, для последовательности кадров получаем набор смещений вдоль осей координат $\{(\Delta X_i, \Delta Y_i) : i \in [2, N]\}$. Сдвиги после удаления дрожания $\{(\Delta \tilde{X}_i, \Delta \tilde{Y}_i) : i \in [2, N]\}$ получаем, применяя к набору $\{(\Delta X_i, \Delta Y_i) : i \in [2, N]\}$ фильтр усреднения.

$$h[n] = \begin{cases} 1/L, & n = \overline{0, L-1} \\ 0 & \end{cases} \quad (2)$$

Подавление дрожания осуществляется смещением каждого кадра на вектор $(\Delta \mathcal{X}_i, \Delta \mathcal{Y}_i)$.

$$(\Delta \mathcal{X}_i, \Delta \mathcal{Y}_i) = (\Delta \tilde{X}_i - \Delta X_i, \Delta \tilde{Y}_i - \Delta Y_i). \quad (3)$$

Алгоритм работает очень быстро, но позволяет получить приемлемый результат на узком классе видеопоследовательностей, так как не оценивает поворот и увеличение.

2.1.2 Восстановление движения камеры

В данном алгоритме [2] производится определение параметров движения камеры в пространстве по известным векторам движения. Каждой точке $p = (x, y)$ на кадре соответствует определенная точка сцены $P = (X, Y, Z)$, и они связаны соотношением (4).

$$(x, y) = (X/Z f_c, Y/Z f_c) \quad (4)$$

f_c – фокусное расстояние камеры.

Движение камеры определяется сдвигами (T_x, T_y, T_z) вдоль соответствующих координатных осей и поворотами $(\Omega_x, \Omega_y, \Omega_z)$ вокруг соответствующих координатных осей. А связь между движением кадра и движением камеры задает равенство (5).

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} -f_c \left(\frac{T_x}{Z} + \Omega_y \right) + x \frac{T_z}{Z} + y \Omega_z - x^2 \frac{\Omega_y}{f_c} + xy \frac{\Omega_x}{f_c} \\ -f_c \left(\frac{T_y}{Z} + \Omega_x \right) + y \frac{T_z}{Z} - x \Omega_z - xy \frac{\Omega_y}{f_c} + y^2 \frac{\Omega_x}{f_c} \end{bmatrix} \quad (5)$$

(u, v) и (x, y) – координаты точки в текущем и предыдущем кадре;

Z – удаленность точки, представленной в предыдущем кадре координатами (x, y) .

Исходя из предположения, что на любом изображении есть область, которой в сцене соответствует плоскость, получаем (6). По утверждению авторов, качественно оценить все параметры нельзя, поэтому определяются только сдвиги вдоль осей (T_x, T_y, T_z) и поворот

вокруг оси Z (Ω_z). Полученная траектория камеры сглаживается, и производится сдвиг кадров соответственно новой траектории камеры.

$$\begin{aligned}
 a &= -f_c \alpha T_X - f_c \Omega_Y & b &= \alpha T_Z - f_c \beta T_X \\
 c &= \Omega_z - f_c \gamma T_X & d &= -f_c \alpha T_Y + f \Omega_X \\
 e &= -\Omega_z - f_c \beta T_Y & f &= \alpha T_Z - f_c \gamma T_Y \\
 g &= -\Omega_z f_c^{-1} + \beta T_Z & h &= \Omega_X f_c^{-1} + \gamma T_Z
 \end{aligned} \tag{6}$$

Основной недостаток алгоритма в том, что точность определения параметров движения не высока. Хотя, значительно выше чем у алгоритма изложенного в предыдущем пункте (2.1.1)

2.2 Подавление дрожания в видео с использованием Feature Tracking

Работу алгоритмов Feature Tracking [3] можно разделить на два этапа:

1) Выбор особых точек (**Selection**).

На этом этапе производится отбор точек, обладающих набором определенных свойств. Вообще говоря, разные методы (например, KLT и IPAN) дают различные наборы особых точек. Но их задача в том, чтобы выбрать точки, окрестность которых значительно отличается от окрестностей близлежащих точек.

2) Определение смещения особой точки между соседними кадрами (**Tracking**).

Этот этап выполняется аналогично алгоритму Motion Estimation из пункта 2.1.

2.2.1 2D Ridge Motion

В алгоритме [1] используется небольшое число особых точек. Проведя сопровождение особых точек, получаем для каждой из них траекторию движения – $x(t)$ и $y(t)$. Движение камеры считается линейным, поэтому траектории приближаются наиболее близкой прямой:

$$\min_{a,b} \sum_{p=1}^N (p(t) - a \cdot t - b) \tag{7}$$

$p(t)$ – точки траектории $x(t)$ или $y(t)$.

Вектор сдвига находится как медиана от векторов сдвига для особых точек. Далее алгоритм работает аналогично подходу, описанному в пункте 2.1.1. И обладает теми же недостатками.

2.3 Подавление дрожания в видео с использованием Image-Based Rendering

На первом этапе [4] определяются и сглаживаются траектории движения отобранных особых точек аналогично алгоритму, описанному в пункте 2.2.1.

Далее для каждой точки, координаты которой получены сглаживанием, определяется значение яркости. Для определения нового значения яркости выбирается несколько кадров,

в которых положение рассматриваемой особой точки близко к полученному сглаживанием (Рис. 1). Яркость в точке вычисляется по формуле (8).

$$c = \sum_i \left(1 - \frac{d_i}{d_{\max}}\right) \cdot c_i, \quad d_{\max} = \max_i(d_i) \quad (8)$$

d_i – расстояние между точками P и P_i ;

c, c_i – цвет точек P и P_i соответственно.

По найденным точкам строится треугольная сетка, и производится визуализация (rendering). Пример построения сетки приведен на рисунке (Рис. 2).

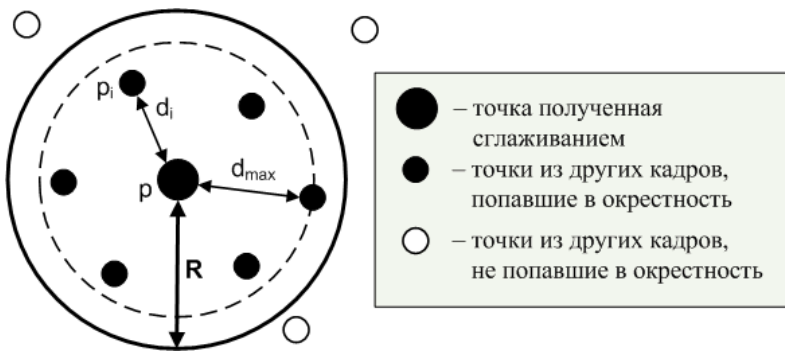


Рис. 1 Определение яркости в точке

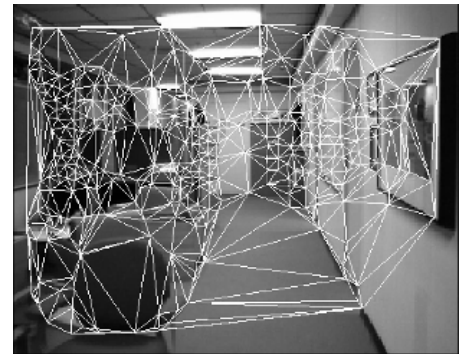


Рис. 2 Пример построения сетки

2.4 Подавление дрожания в видео с анализом кадра целиком

2.4.1 Алгоритм подавления дрожания кадра, строящий аффинную модель движения [5]

Изменение положения кадра оценивается с помощью аффинной модели.

$$\begin{pmatrix} x^n \\ y^n \end{pmatrix} = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \cdot \begin{pmatrix} x^{n-1} \\ y^{n-1} \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = A_n \cdot \begin{pmatrix} x^{n-1} \\ y^{n-1} \end{pmatrix} + b_n \quad (9)$$

x^n, y^n – координаты точки в n-ом кадре.

Таким образом, производится оценка поворота, увеличения и сдвига. Параметры A_n и b_n определяются при помощи алгоритма градиентного спуска с функцией цены (10).

$$\sum_x \left((I_n(x) - I_{n-1}(A_n \cdot x + b_n))^2 + \beta \right)^{1/2} \quad \beta > 0 \quad (10)$$

После определения матриц A и векторов b для всех кадров проводится их покомпонентное сглаживание (например, с помощью фильтра Кальмана). В результате, каждому A_n и b_n ставятся в соответствие \tilde{A}_n и \tilde{b}_n , которые отражают истинное движение кадра (без дрожания). С использованием \tilde{A}_n и \tilde{b}_n строится стабилизированное видео, в котором отсутствует дрожание кадра.

Данный алгоритм позволяет получить наилучший результат в сравнении с остальными рассмотренными в обзоре алгоритмами, но требует значительных вычислительных затрат.

3 Предлагаемый алгоритм

Предлагаемый алгоритм подавления эффекта дрожания состоит из трех шагов:

- 1) определение смещения между соседними кадрам;
- 2) удаление дрожания;
- 3) заполнение неизвестных областей на краях кадра.

3.1 Определение смещения между соседними кадрами

Определение глобального смещения между соседними кадрами выполняется на основе информации о локальном движении при помощи алгоритма, описанного авторами в статье [6]. Особенность подхода заключается в том, что алгоритм не ориентирован на какой-либо определенный алгоритм определения локального движения (Motion Estimation, Feature Tracking, Optical Flow), а может работать с результатом любого из них.

В результате этого шага для каждого кадра I_n определяются параметры смещения: x_n, y_n - сдвиги вдоль осей, φ_n - угол поворота, z_n - увеличение. Иными словами, оценивается траектория движения камеры (используется прощенная модель движения).

3.2 Удаление дрожания

Найденные траектории движения сглаживаются покомпонентно с помощью низкочастотной фильтрации. После чего кадры сдвигаются согласно разнице между исходной и сглаженной траекториями. Сдвиг вычисляется по формуле (11).

$$\mathcal{F} = T^{-1} \cdot \tilde{T} \quad (11)$$

T – смещение кадра в исходной видеопоследовательности;

\tilde{T} – смещение кадра после сглаживания;

\mathcal{F} – сдвиг, компенсирующий дрожание.

3.3 Заполнение неизвестных областей на краях кадра

Поскольку кадры сдвигаются относительно исходного положения, на их краях появляются неизвестные области. Самый простой способ решения этой проблемы заключается в обрезании краев кадра. Но такой подход не очень хорош, так как в результате теряется часть информации.

Для заполнения неизвестных областей авторы предлагают использовать панорамное изображение, реконструируемое в процессе работы алгоритма по известным параметрам глобального смещения между соседними кадрами (Рис. 3, Рис. 4). Такой подход не требует сложных дополнительных вычислений и обеспечивает приемлемое визуальное качество.

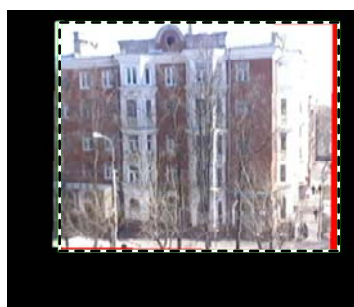


Рис. 3 Результат без заполнения краев

Рис. 4 Результат с заполнением краев

Результат работы предложенного алгоритма представлен на рисунке (Рис. 6).

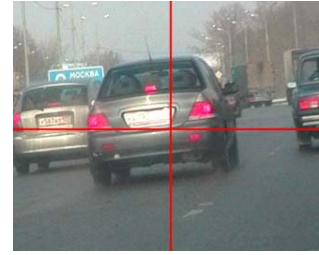


Рис. 5 Последовательность соседних кадров из видео с дрожанием



Рис. 6 Последовательность соседних кадров из видео после подавления дрожания

4 Определение наличия дрожания

Для определения наличия дрожания на некотором фрагменте видеопоследовательности анализируется соответствующий участок траектории движения независимо по каждой компоненте. Для каждой компоненты траектории анализируется ее спектр.

Как показали исследования, амплитуды спектра можно разделить на три класса: «форма», «дрожание» и «шум» (Рис. 7).

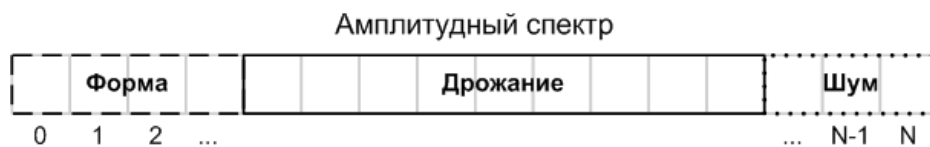


Рис. 7 Разделение амплитуд спектра на классы

Для определения наличия дрожания анализируются амплитуды из класса «дрожание». На видеопоследовательностях без дрожания все амплитуды из этого класса лежат ниже некоторого порога (Рис. 8). Таким образом, если на некотором фрагменте траектории хотя бы одна амплитуда из класса «дрожание» превышает установленный порог, то считается, что на соответствующем фрагменте видео присутствует дрожание.

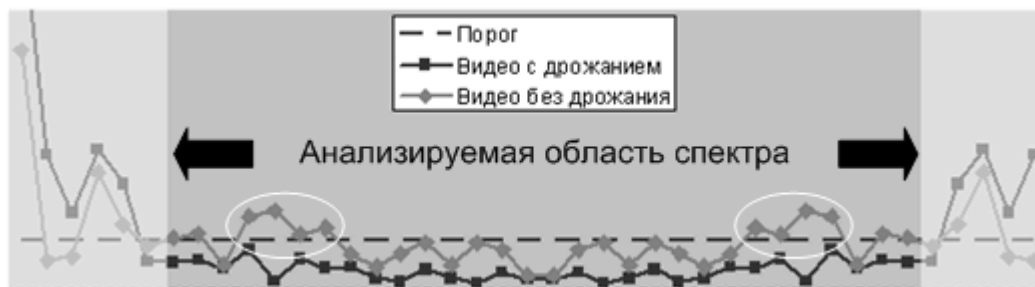


Рис. 8 Пример определения наличия дрожания

Предложенный алгоритм дает возможность не выполнять подавление дрожания на участках видео с плавным движением, что позволяет повысить скорость обработки видеопоследовательностей и снизить количество артефактов, вносимых алгоритмом подавления эффекта дрожания.

5 Заключение

В статье проанализированы наиболее известные алгоритмы подавления дрожания в видео и предложен собственный алгоритм, учитывающий их достоинства и недостатки. Предлагаемый алгоритм применим для обработки любых видеопоследовательностей, так как он анализирует видео на предмет наличия дрожания и обрабатывает только видео в которых эффект дрожания присутствует. Алгоритм определяет сдвиг, поворот и увеличение с высокой степенью точности и обрабатывает 30 кадров размером 352x288 пикселей в секунду на Pentium-4 2.8 ГГц за счет использования алгоритма быстрого и надежного определение глобального движения в видеопоследовательностях [6]. Особо стоит отметить, что предложенный алгоритм позволяет заполнять неизвестные области появляющиеся на краях кадров в результате компенсации дрожания.

6 Список литературы

- [1] Kwong Kin, Wah Eric, «Video Stabilization Algorithms», ELEC 533 Project Report, 2003
- [2] Irani M., Rousso B., Peleg S., «Recovery of Ego-Motion Using Image Stabilization», CVPR '94, pp. 454-460, 1994
- [3] Tomasi C., Kanade T., «Detection and Tracking of Point Features», Carnegie Mellon University Technical Report CMU-CS-91-132, April 1991
- [4] Buehler C., Bosse M., McMillan L., «Non-Metric Image-Based Rendering for Video Stabilization», Proceedings of CVPR, 2001
- [5] Litvin A., Konrad J., Karl W.C., «Probabilistic video stabilization using Kalman filtering and mosaicking», Symposium on Electronic Imaging, Image and Video Communications, 2003
- [6] Солдатов С.А., Стрельников К.Н., Ватолин Д.С., «Быстрое и надежное определение глобального движения в видеопоследовательностях», Труды конференции Graphicon-2006, стр. 430-437, Новосибирск, Академгородок, Россия, Июль 2006