

ГИБРИДНЫЙ ПОДХОД К РЕШЕНИЮ ЗАДАЧИ ОЦЕНИВАНИЯ ГЕОМЕТРИЧЕСКИХ ИСКАЖЕНИЙ КАДРОВ ВИДЕОПОСЛЕДОВАТЕЛЬНОСТЕЙ

Слынько Ю. В., Лагуткин В. Н., Лукьянов А. П.

ОАО МАК "Вымпел"

1 ВВЕДЕНИЕ И ПОСТАНОВКА ЗАДАЧИ

Одной из базовых проблем в обработке видеопоследовательностей является задача определения параметров геометрических преобразований кадров. К решению этой задачи сводятся такие задачи, как стабилизация изображений, сопровождение движущихся объектов и оценка параметров их движения, создание мозаики и др.

Более того, многие алгоритмы обработки изображений, такие как выделение переднего плана, распознавание образов и другие, требуют для своей работы предварительной стабилизации изображений.

Существуют различные методы оценки параметров деформации кадра: на основе оптического потока внутри кадра [1], на основе выделения характерных точек изображения [2], напрямую использующие яркости точек входного изображения [3].

Однако все эти алгоритмы не рассчитаны на работу в реальном времени на неспециализированных вычислительных средствах. Они работают устойчиво, только если входные изображения удовлетворяют выбранной математической модели, и зачастую не обладают достаточной гибкостью при изменении этой модели или несоответствии ей входных изображений.

В данной работе предложен метод решения такой задачи. Причем основными свойствами предложенных алгоритмов является высокая скорость работы, устойчивость и надежность в различных ситуациях.

В качестве моделей искажений в предложенном методе могут быть использованы различные модели: аффинные, квадратичные, проективные преобразования и др.

При разработке алгоритма особенно учитывалось то, что при работе с реальными изображениями возникают различные отклонения от идеальной математической модели: смазы, нелинейные искажения в оптике камеры, потеря информации при кодировании видео и целый ряд других.

В качестве ключевых особенностей предложенного алгоритма можно выделить две:

1. Алгоритм основан на объединении трех основных подходов к решению данной задачи, что позволяет компенсировать их недостатки и подчеркнуть достоинства.

2. В алгоритмах определения сдвига используется детальный анализ функции корреляции, что позволяет оценивать точность и надежность принятия решения и ограничить вычисления тогда, когда эти показатели достигнут заданных значений.

2 ОБЩИЙ ВИД АЛГОРИТМА

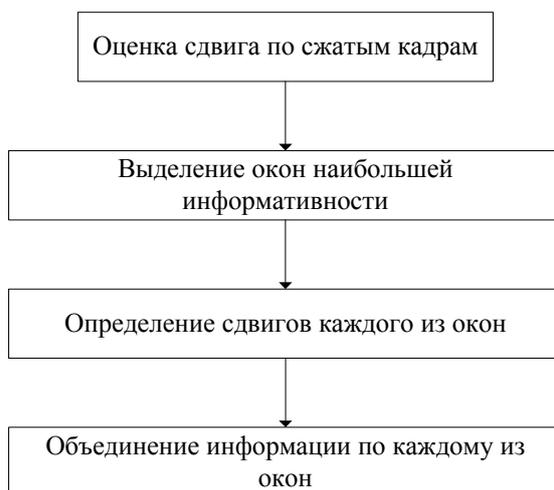


Рис. 1. Принципиальная блок-схема алгоритма.

Блок-схема алгоритма представлена на рис. 1.

В качестве первого шага на пути к решению задачи оценки геометрических искажений используется оценка сдвига с использованием пирамиды сжатых кадров. При этом вычисления не стоит вести до достижения масштаба исходного кадра, а следует прервать раньше, что может существенно уменьшить количество вычислений. Метод поиска сдвига описан в следующем разделе. Поиск по сжатым кадрам позволяет определить большие сдвиги, вплоть до полукадра, при этом обеспечивая хорошую устойчивость и надежность.

Затем производится поиск наиболее информативных областей кадра. Эти области выбираются из критерия максимума автокорреляционной функций, причем так, чтобы они находились по возможности дальше друг от друга. Это в целом сходно с нахождением характерных точек на изображении, но в данной работе такие области (корреляционные окна) не являются точечными, а состоят из достаточного количества точек (больше 100). Затем определяется сдвиг каждого из этих окон (часть 3).

Такой метод позволяет, как и в методе характерных точек, использовать только наиболее существенную информацию. Но сдвиг каждой точки (окна) определяется непосредственно и задача нахождения соответствия между точками на последовательных кадрах не стоит, что очень существенно при больших межкадровых сдвигах. Также данный метод позволяет работать с размытыми изображениями или изображениями, на которых мало характерных точек.

После определения сдвига по каждой из областей, происходит объединение информации и нахождение общего преобразования всего кадра (часть 4). Здесь используются методы, схожие с методами определения параметров преобразования на основе оптического потока. Ведь набор векторов сдвигов каждого из окон можно считать оптическим потоком кадра. Однако, в отличие от оптического потока, каждый вектор здесь вычислен гораздо надежнее, поскольку вычисление основано на гораздо большем количестве информации. Как следует из метода определения сдвига, помимо собственно сдвигов каждого из окон были определены ковариационные матрицы ошибок оценки эти векторов, что также можно использовать при определении полного вектора параметров трансформации.

3 ОПРЕДЕЛЕНИЕ СДВИГА

Для определения сдвига используется метод поиска минимума функции $X_{(F^1, F^2)}(dx, dy)$ среднеквадратичной невязки (корреляции) двух кадров F^1 и F^2 . Вычисления ведутся по области кадра $M(dx, dy)$ площадью $S(M)$:

$$X_{(F^1, F^2)}(dx, dy) = \frac{1}{S(M)} \sum_{(i,j) \in M(dx, dy)} (F^1_{ij} - F^2_{i+dx, j+dy})^2. \quad (1)$$

Искомые параметры $(d\hat{x}, d\hat{y})$ ищутся в области D :

$$(d\hat{x}, d\hat{y}) = \arg \min_{(dx, dy) \in D} X_{(F^1, F^2)}(dx, dy). \quad (2)$$

Для нахождения минимума данной функции различные методы спуска неприемлемы, т.к. часто приводят к нахождению побочного локального минимума. Для устойчивой работы необходимо применять методы перебора. Таким образом, находится целочисленный сдвиг. Что касается дробной части сдвига, то различные методы ее нахождения описаны, например, здесь [4].

Для ускорения перебора, улучшения надежности нахождения минимума и уменьшения ошибки определения сдвига предлагается проводить подробный анализ функции корреляции.

Предлагается следующая схема. Пусть у нас вычислены значения функции $X_{(F^1, F^2)}(dx, dy)$ в некоторой области D . Искомые параметры соответствуют минимуму функции, поэтому для оценки качества их определения необходимо исследовать поведение функции $X_{(F^1, F^2)}(dx, dy)$ в районе ее минимума.

Доверительная область, в которой могут лежать значения функции $X_{(F^1, F^2)}(dx, dy)$, обусловленные шумом, будет следующая:

$$\tilde{D} = \{dx_i, dy_i : X_{(F^1, F^2)}(dx_i, dy_i) < m_i + B \cdot \sigma_i\}, \text{ где} \quad (3)$$

$$m_i = \frac{X_{\min}}{N_p} \cdot N_i, \quad \sigma_i = \frac{X_{\min}}{N_p} \cdot \sqrt{2N_i},$$

$$X_{\min} = \min_{(dx, dy) \in D} X_{(F^1, F^2)}(dx, dy), \quad N_p = S(M(d\hat{x}, d\hat{y})), \quad N_i = S(M(dx_i, dy_i)),$$

B – порог, определяющий доверительную вероятность ошибки.

Таким образом, для точки минимума функции $X_{(F^1, F^2)}(dx_i, dy_i)$ мы получили область, в которой с заданной вероятностью находится истинное решение. Эта область достаточно полно характеризует качество определения сдвига.

Зная такую доверительную область можно существенно ускорить работу алгоритма без потери надежности. В частности, при вычислении сдвига пирамидой по сжатым кадрам на каждом шаге, кроме первого, можно производить поиск минимума не по всей области возможных сдвигов D , а по области \tilde{D} , вычисленной на предыдущем шаге, что может давать прирост в скорости на несколько порядков.

Также можно использовать не полный перебор, а прореженный. Затем по прореженному множеству значений функции $X_{(F^1, F^2)}(dx, dy)$ строится область \tilde{D} , внутри которой производится полный перебор.

При таком подходе надежность определения точки минимума не меньше, чем при полном переборе, а скорость работы значительно выше.

Однако оценка (3) является существенно нелинейной и зачастую неудобной для анализа. Ее можно заменить более простой оценкой, ограничившись вторыми моментами распределения, то есть использовать гауссову модель ошибок с ковариационной матрицей оценивания искомой величины (в данном случае - сдвига). Обратная ковариационная матрица ошибок оценивания сдвигов

$$C^{-1} = \begin{vmatrix} A & B \\ B & C \end{vmatrix} \quad (4)$$

имеет размерность 2×2 , а ее элементы представляют собой оценки параметров эллипса, аппроксимирующего область \tilde{D} методом наименьших квадратов в виде:

$$A(x - d\hat{x})^2 + 2B(x - d\hat{x})(y - d\hat{y}) + C(y - d\hat{y})^2 = 0. \quad (5)$$

4 ОПРЕДЕЛЕНИЕ ПАРАМЕТРОВ ТРАНСФОРМАЦИИ ВСЕГО КАДРА

После оценки сдвига каждого из выбранных окон необходимо вычислить параметры трансформации всего кадра. Здесь пригодны все те подходы, которые используются в методах с использованием оптического потока и характерных точек.

Принципиально решение данной задачи сводится к нахождению минимума:

$$\hat{Q} = \arg \min_Q \sum_{k=1}^N (\bar{\rho}_k - \tau_Q(\bar{r}_k)) C_k^{-1} (\bar{\rho}_k - \tau_Q(\bar{r}_k))^T, \text{ где} \quad (6)$$

\bar{r}_k – координаты центров окон в кадре $F(t-I)$, $\bar{\rho}_k$ – координаты тех же областей в кадре $F(t)$ (вычисленные по формулам (1), (2)), N – количество прямоугольников, C^{-1} – ковариационная матрица ошибок, $\tau_Q(\bar{r}_k)$ – функция преобразования координат, определяемая выбранной моделью.

Однако в данном случае имеются существенные преимущества перед стандартными подходами:

1. Если решать задачу минимизации фильтром Калмана, то, зная ошибку определения каждого измерения (4), можно остановить вычисления, когда точность суммарной оценки достигнет желаемой величины.

2. При введении регуляризации в решение (6), с учетом ковариационной матрицы (4), достигается значительное увеличение качества работы в трудных случаях, например, при наблюдении линейной картины.

5 РЕЗУЛЬТАТЫ И ВЫВОДЫ

Исследование работы алгоритмов проводилось на ПК со следующими характеристиками: процессор - Intel® Pentium® 4 2.4ГГц, память - 512Мб, видеокарта - NVIDIA GeForce4® MX 440, операционная система - Microsoft® Windows™ 2000.

Сводные результаты работы представлены в табл. 1.

Тестирование точностных характеристик и скорости проводилось на синтетических последовательностях. Также предлагаемый алгоритм был протестирован не менее чем на ста реальных последовательностях различного характера. Однако, дать численную оценку качества работы на таких последовательностях трудно. В качестве такой оценки была выбрана вероятность сбоя (столбец 7 табл. 1.). Также был проведен анализ работы всего комплекса алгоритмов обработки изображений, включающего в себя как представленные алгоритмы, так и чтение последовательности с диска и отображение результатов на экран в условиях реального времени (последний столбец табл. 1.).

Точность определения сдвига, пиксели	Максимальный сдвиг, % размера кадра	Точность определения угла, °	Максимальный угол, °	Точность определения коэффициента масштабирования, %	Количество кадров (320x240) в секунду	Вероятность сбоя на реальных последовательностях, %	Максимальный размер кадра, обрабатываемый в реальном времени
0.03	50	0.034	10	0.2	715	<0.1	1600x900

Табл. 1.

Было показано, что представленные алгоритмы имеют высокую точность оценки параметров – на уровне других известных алгоритмов. При этом такие параметры, как максимальный оцениваемый сдвиг, скорость работы и устойчивость в "трудных" условиях превосходят большинство известных методов.

Также следует отметить, что разработанные алгоритмы обладают большой гибкостью: позволяют работать с различными моделями геометрических преобразований, исключать влияние движущихся объектов, устойчивы к различным моделям шумов и т.д.

Литература

1. Negahdaripour S. and Lee S. // "Motion Recovery from Image Sequences Using First-Order Optical Flow Information," IEEE Workshop Visual Motion, Princeton, N.J., 1991. Oct. pp. 132–139.

2. Lustman F., Faugeras O.D., and Toscani G. // "Motion and Structure from Motion from Point and Line Matching." Proc. First Int'l Conf. Computer Vision, London, 1987. pp. 25–34.

3. Horn B. K. P. and Weldon E. J. Jr. // Direct methods for recovering motion. International Journal of Computer Vision, 1988. 2: 51–76.

4. Ким А.К., Колесса А.Е., Лагуткин В.Н., Лотоцкий А.В., Репин В.Г. // Алгоритмы идентификации и подавления нестационарного мешающего фона и повышение разрешающей способности в оптическом сенсоре с хаотически колеблющейся оптической осью и динамическими аберрациями. Радиотехника, 1998. № 12.



HYBRID APPROACH TO THE SOLUTION OF THE PROBLEM OF WARPING ESTIMATION OF FRAMES IN VIDEO SEQUENCES

Slynko Y., Lukyanov A., Lagutkin V.

ОАО МАК "Vimpel"

In the report problems of estimation of warping parameters of frames in video sequence, image stabilizing and object tracking are concerned. Affine, projective and quadratic warping are used as a model of warping.

Proposed algorithms were created to provide two main qualities: robust operation in a wide range of shooting conditions and working in a real time on conventional computational systems.

A combination of three main approaches is used to solve this problem: method of the optical flow, method of feature points and direct method using brightness of points. That allows us to use all their advantages and compensate drawbacks.

As a great improving of these algorithms a method of detailed analysis of correlation function of frames is used. That allows avoiding unnecessary computations, control accuracy of estimation and confident region of values of these estimation, what finally result in increasing in accuracy, robustness and speed of calculations.

Proposed algorithms were tested on synthetic sequences. The resulting accuracy of estimation of shift is up to 0.07 pixels, estimation of angle is up to 0.034° with the processing speed up to 715 frames per sec for the frame size of 320x240 pixels.

Efficiency of algorithms was tested on more then hundred real videos of different character and quality. A crash probability was estimated and amounted less then 0.1%.

One should note the proposed algorithms are agile and work with different warping models, able to exclude moving objects, robust with different types of noise etc.

