

УДК 004.932

СРАВНЕНИЕ ДВУХ СТРУКТУРНЫХ АЛГОРИТМОВ ОЦЕНКИ ПАРАМЕТРОВ ГЕОМЕТРИЧЕСКИХ ПРЕОБРАЗОВАНИЙ ИЗОБРАЖЕНИЙ

Стротов В.В., к.т.н., доцент кафедры автоматизи и информационных технологий в управлении Рязанского государственного радиотехнического университета, e-mail: aitu@rsreu.ru.

Ключевые слова: геометрические преобразования, структурный алгоритм, оценка параметров, опорные участки изображения.

Предварительные замечания

Оценка параметров геометрических преобразований изображений является одной из важных задач, решаемых в системах обработки видеоизображений. Получаемые результаты используются при решении таких задач, как выделении движущихся объектов, оценка их координат, а также для электронной стабилизации изображений [1-3]. При этом, как правило, оценку параметров преобразований необходимо производить в темпе поступления информации, что накладывает значительные ограничения на вычислительную сложность используемых алгоритмов.

В современной литературе предложено достаточное количество методов решения данной задачи [1, 4]. Их можно разделить на несколько больших классов:

- методы морфологического анализа изображений;
- методы прямого сопоставления изображений;
- методы анализа оптического потока;
- структурные методы.

Внутри каждого из представленных классов существует несколько направлений развития методов оценки геометрических деформаций. Однако можно утверждать, что для вычислительных систем реального времени, построенных на основе процессоров общего назначения и цифровых сигнальных процессоров, первые три класса методов, как правило, являются вычислительно сложными. Следовательно, наиболее пригодными для таких систем представляются структурные методы.

В данной работе автор ставит перед собой цель на основе результатов сравнения наиболее характерных алгоритмов различных типов, относящихся к классу структурных, сформулировать рекомендации относительно их применения в системах обработки видеоизображений реального времени.

Структурные методы оценки параметров геометрических преобразований

Данный класс объединяет подходы, которые для оценки параметров геометрических преобразований используют информацию о положении так называемых структурных элементов на обрабатываемых изображениях. Фактически, используется более компактное описание изображений по сравнению с традиционным яркостным представлением. Это позволяет уменьшить вычислительные затраты на сопоставление изображений.

В работе рассмотрены два структурных алгоритма оценки параметров геометрических преобразований изображений: алгоритм, основанный на выборе и оценке положения опорных участков изображений, и алгоритм SIFT, основанный на оценке положения особых точек изображений. Представлены результаты сравнительных экспериментальных исследований. Даны рекомендации по использованию данных алгоритмов в системах обработки видеоизображений реального времени.

Структурные методы оценки преобразований можно разделить на несколько групп по типу структурного элемента. Основными структурными элементами являются [2]:

- участки изображений;
- особые точки;
- отрезки линий, окружностей или других параметрических фигур.

Анализ литературы показывает, что использование отрезков линий и других параметрических фигур в качестве структурных элементов в системах реального времени возможно только при условии соответствия обрабатываемых изображений довольно жёстким требованиям. Это связано с тем, что, во-первых, наличие достаточного количества стабильных элементов такого вида сложно гарантировать, а, во-вторых, этап выделения структурных элементов для данного типа имеет чрезмерную вычислительную сложность. Следовательно, в системах, осуществляющих обработку изображений в реальном времени, предпочтительно использовать структурные алгоритмы на основе оценки положения небольших участков и особых точек изображений.

Алгоритмы оценки параметров некоторых видов преобразований, таких как сдвиг изображений или его частей на основе оценки положения опорных участков изображений достаточно хорошо разработаны и освещены в литературе. В частности, определение межкадровых искажений сцены на основе слежения за участками размером 4 x 4 пикселя применяется в алгоритмах видеокомпрессии MPEG. При этом рассматриваемые алгоритмы изначально не были ориентированы на обеспечение высокой точности при работе в системах обработки видеoinформации реального времени из-за отсутствия процедур выбора опорных участков. Более поздние модификации, например предложенные в [5], предлагают эффективные критерии выбора. Это позволяет снизить количество опорных участков до 5-15 в зависимости от особенностей решаемой задачи, что приводит к значительному увеличению производительности с сохранением необходимой точности.

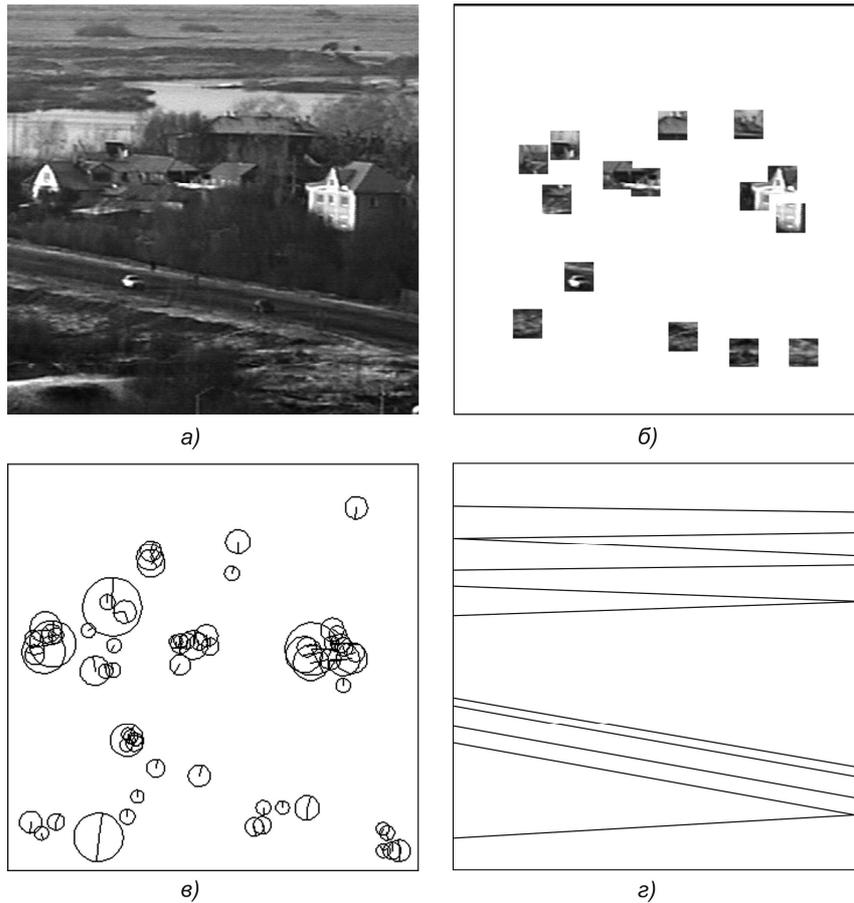


Рис. 1. – Различные формы представления изображения: а) – яркостное представление (400x400 точек, 8 бит/пиксель) б) – в виде набора наиболее информативных участков изображения (15 участков размером 24x24 пикселей), в) – в виде набора особых точек (детектор SIFT, 80 точек), г) – в виде набора линий (детектор на основе преобразования Радона, 11 линий)

Структурные алгоритмы оценки параметров геометрических преобразований, использующие в качестве опорных элементов особые точки различных видов, начали активно развиваться на рубеже двадцатого и двадцать первого веков. На данный момент предложено немало способов для выделения и описания особых точек изображений. Они варьируются по таким показателям как эффективность, устойчивость и быстродействие. Нередко данные показатели вступают в противоречие, поэтому выбор подходящего алгоритма для реализации в конкретной аппаратуре реального времени всегда является предметом для исследования. Во многих работах, таких как [6] показано, что наиболее точным является известный алгоритм SIFT, предложенный в 1999 году. Более поздние разработки, такие как алгоритмы PCA-SIFT [6], SURF [7], ORB [8] и их модификации отличаются значительным увеличением скорости работы при некотором снижении точности и устойчивости, а также снижением инвариантности к некоторым видам геометрических преобразований.

Алгоритм оценки параметров геометрических преобразований изображений на основе оценки положения опорных участков

1. Рассматриваемый алгоритм был достаточно подробно рассмотрен в работе [5]. Он позволяет оценивать параметры преобразований евклидовой группы и мас-

штаба на изображениях, наблюдаемых как в одном, так и в нескольких спектральных диапазонах. Принцип работы данного алгоритма состоит в следующем.

На исходном изображении $l(i, j)$ выбираются опорные участки количеством не более 15 и размером не более 32×32 пикселей. Положение опорных участков $(\hat{\alpha}_m, \hat{\beta}_m)$ определяется по критерию:

$$(\hat{\alpha}_m, \hat{\beta}_m) = \arg \min_{\alpha_m, \beta_m} \left(\frac{1}{\sum_{(i,j) \in S_m} (\Delta_i l(i,j))^2} + \frac{1}{\sum_{(i,j) \in S_m} (\Delta_j l(i,j))^2} \right),$$

где $m = \overline{1, M}$, $\Delta_i l(i, j)$ означает операцию численно-го дифференцирования изображения $l(i, j)$ по координате i , а S_m – множество точек участка. Размер и количество участков определяются параметрами аппаратной части системы обработки видеoinформации и требованиями по достижимой точности.

2. При поступлении изображения очередного кадра на нём определяются положения всех опорных участков с субпиксельной точностью путём нахождения положения точки экстремума выбранной критериальной функции.

$$(\hat{i}_m(n), \hat{j}_m(n)) = \arg \min_{i, j \in W_m(n)} (F_{\text{крит}}(i, j)), \text{ где}$$

$$F_{\text{крит}}(i, j) = \sum_{\alpha=1}^{a_m} \sum_{\beta=1}^{b_m} \left| I \left(i + \alpha - \frac{a_m}{2}, j + \beta - \frac{b_m}{2} \right) - w_m(\alpha, \beta, n) \right|.$$

Здесь $W_m(n)$ – зона поиска эталона w_m в n -м кадре. Повышение точности достигается применением полиномиальной интерполяции $F_{\text{крит}}(i, j)$ в области целочисленного экстремума.

3. Из псевдорешения системы линейных уравнений, описывающей связь координат опорных участков в текущем и предыдущем кадрах, находятся предварительные оценки параметров межкадрового преобразования согласно заданной модели.

4. Так как на наблюдаемом изображении может находиться небольшое количество движущихся относительно фона объектов, которые могут быть выбраны в качестве опорных, возможно возникновение ошибок определения параметров преобразований. Для исключения таких ситуаций вычисляется вектор невязки системы уравнений. На основе анализа элементов данного вектора из рассмотрения исключаются некоторые опорные участки и определяются уточнённые оценки параметров межкадрового преобразования.

5. На основе найденных на предыдущем этапе оценок определяются параметры геометрического преобразования.

6. Производится замена опорных участков расположенных близко к границам изображения или имеющих большую невязку, после чего ожидается поступления изображения следующего кадра.

Алгоритм оценки параметров геометрических преобразований изображений на основе оценки положения особых точек

Алгоритм SIFT является достаточно известным и хорошо описан, например, в работе [9]. В данной работе коротко остановимся на основных его этапах.

1. На изображениях с помощью выбранного детектора выделяются особые точки.

2. Для каждой выделенной точки находится некоторый набор данных, инвариантный к оцениваемым преобразованиям – дескриптор. Он характеризует окрестность особой точки.

3. По совпадению дескрипторов на исходном и трансформированном изображениях выделяются соответствующие друг другу особые точки.

4. На основе набора совпавших особых точек производится оценка параметров преобразования изображений согласно выбранной модели преобразования.

Основным моментом в детектировании особых точек является построение пирамиды гауссианов (Gaussian) и разностей гауссианов (Difference of Gaussian, DoG). Разностью гауссианов называют изображение, полученное путем поэлементного вычитания одного гауссиана исходного изображения из гауссиана с другим радиусом размытия (1).

$$D(x, y, \sigma) = (G(x, y, \sigma_s) - G(x, y, \sigma_{s+1})) * I(x, y). \quad (1)$$

σ – радиус размытия, G – гауссово ядро, $*$ – операция свертки. В данном алгоритме размытие изображения гауссовым фильтром считается эквивалентным масштабированию изображения, причём определённая степень размытия соотносится с конкретным масштабом.

После построения пирамид находятся точки, являющиеся локальными экстремумами разности гауссианов. Для этого каждая точка текущего изображения DoG сравнивается с её восьмью соседями и с девятью соседями в DoG, находящихся на уровень выше и ниже в пирамиде. Если этот отклик в данной точке больше (меньше) всех соседей, то она принимается за точку локального экстремума. Далее производится субпиксельное уточнение координат каждой особой точки, а затем исключаются точки с малым откликом и точки, лежащие на границах.

Дополнительно для каждой особой точки производится оценка величины, называемой ориентацией. Она вычисляется исходя из направлений градиентов точек, лежащих в окрестности особой на изображении в пирамиде гауссианов, с масштабом наиболее близким к масштабу точки.

Далее вычисляется дескриптор особой точки, который в методе SIFT является вектором. Как и направление опорной точки, дескриптор вычисляется на гауссиане, ближайшем по масштабу к опорной точке на основе градиентов в некотором окне, окружающем опорную точку. Перед вычислением дескриптора это окно поворачивают на угол, определяемый направлением опорной точки, чем и достигается инвариантность относительно поворота. Затем вычисляются гистограммы градиентов в окрестности по следующим правилам:

1) каждая гистограмма покрывает участок в 360 градусов и делит его на 8 частей;

2) в качестве весового коэффициента берется значение гауссова ядра, общего для всего дескриптора;

3) в качестве дополнительных весовых коэффициентов берутся коэффициенты трилинейной интерполяции.

Дескриптор ключевой точки состоит из всех полученных гистограмм. Размер его может варьироваться, однако на практике обычно используются дескрипторы размерности 128 компонент.

Оценки параметров геометрических преобразований вычисляются по схеме, похожей на ту, что применяется в алгоритме, описанном выше. Отличие состоит в том, что количество особых точек обычно на порядок больше, чем количество участков, поэтому для нахождения достаточно точных оценок параметров преобразований необходимо использовать статистические методы согласования полученных экспериментальных данных с выбранной моделью, например RANSAC [10].

Экспериментальные исследования

Сравнительные исследования двух описанных выше алгоритмов производилось на совокупности из двадцати натуральных и искусственных (синтезированных) полутонных видеосюжетов размером от 640x480 до 768x576 пикселей. Преобразования фона на данных видеосюжетах представляют собой совокупность преобразований сдвига, поворота и изменения масштаба изображения. На изображениях присутствуют движущиеся объекты



общей площадью не более 3% от площади кадра. Видеопоследовательности получены от различных датчиков видеoinформации, таких как телевизионные камеры с чересстрочной и прогрессивной развёрткой, тепловизоры. Наблюдались различные виды фоноцелевой обстановки в различных погодных условиях. Всё это позволило получить информацию о поведении исследуемых алгоритмов на видеопоследовательностях различного качества.

Условно разделим все видеопоследовательности на три класса.

1. Видеопоследовательности низкого качества, характеризующиеся низким контрастом, и, как следствие, низким отношением сигнал/шум (менее 8), а также малым количеством неподвижных объектов, выделяющихся на общем фоне.

2. Видеопоследовательности среднего качества, характеризующиеся средним контрастом и отношением сигнал/шум около 8, а также достаточным количеством контрастных неподвижных объектов фона.

3. Видеопоследовательности высокого качества, характеризующиеся отношением сигнал/шум более 8, а также наличием большого количества контрастных неподвижных объектов фона.

Примеры кадров видеопоследовательностей из каждого класса приведены на рис. 2.

Эталонные параметры геометрических преобразований на натуральных сюжетах были оценены человеком-экспертом с достаточной точностью. Сравнение алгоритмов производилось по точности определения параметров сдвига, поворота и изменения масштаба. В работе представлена часть результатов исследований, иллюстрирующих наиболее характерные особенности в работе алгоритмов, остальные результаты учтены при составлении рекомендаций по использованию алгоритмов.



а)



б)



в)

Рис. 2. Примеры кадров видеопоследовательностей, отнесённых к разным классам по качеству: видеопоследовательность низкого качества а), видеопоследовательность среднего качества б) и видеопоследовательность высокого качества в)

На рис. 3 а)-в) приведены результаты исследования точностных параметров структурного алгоритма на основе выбора опорных участков (алгоритм 1) и алгоритма SIFT (алгоритм 2) в зависимости от уровня аддитивного шума. Для каждой видеопоследовательности была произведена оценка контраста и произведено добавление помехи для достижения необходимого соотношения сигнал/шум. Параметры алгоритмов были выбраны следующими: для алгоритма 1 количество опорных участков равно 10, размер участков – 32×32 пикселя; для алгоритма 2 количество выделяемых точек составляло порядка 200. Параметры алгоритмов были выбраны такими для того, чтобы уравнивать время на обработку кадра (для текущих реализаций).

Из графиков, приведённых на рис. 3 а)-в) можно сделать вывод о том, что оба алгоритма работоспособны на данной подборке видеопоследовательностей и обеспечивают приемлемую точность для решения большинства задач в системах обработки видеоизображений. Однако можно утверждать, что алгоритм 1 показал себя менее требовательным к качеству видеопоследовательностей, тогда как алгоритм 2 показал лучшую точность при качестве видео выше среднего.

На рис. 3 г) приведены результаты исследования инвариантности алгоритмов к различным видам преобразований. Исследования показали, что алгоритм 2 инвариантен ко всем видам преобразований, а алгоритм 1 при инвариантности к сдвигу и изменению масштаба недостаточно инвариантен к повороту изображения.

Таким образом, по результатам экспериментов можно рекомендовать использование структурного алгоритма, основанного на выборе и слежении за опорными участками в случае, когда качество наблюдаемых видеопоследовательностей невелико и существует априорная информация о том, что максимальный поворот изображений не будет превышать 45 градусов.

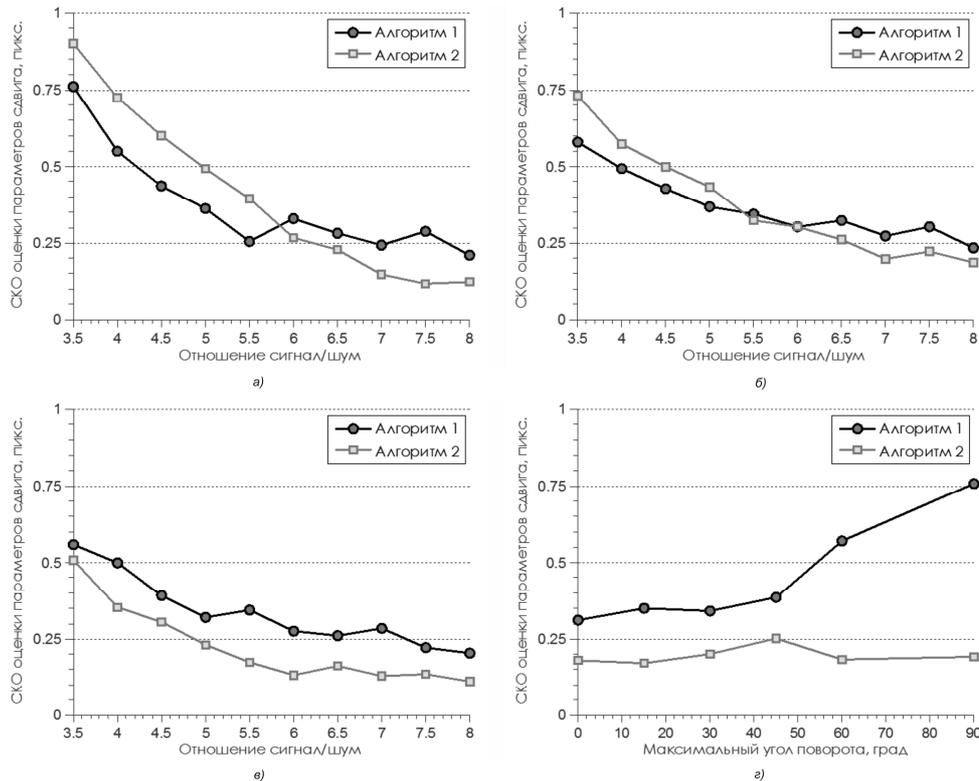


Рис. 3. Результаты исследования алгоритмов: исследования устойчивости к аддитивному шуму, усреднённые по видеопоследовательностям различных классов: низкого качества а), среднего качества б) и высокого качества в); исследования устойчивости к повороту изображения г)

Заключение

В данной работе рассмотрены представители двух классов структурных алгоритмов – алгоритм на основе выбора и слежения за опорными участками и известный алгоритм на основе слежения за особыми точками (SIFT). Проведены экспериментальные исследования, по результатам которых даны рекомендации по использованию алгоритмов в системах обработки видеоизображений.

Исследования проводились при финансовой поддержке Министерства образования и науки Российской Федерации (договор № 16.120.11.300-МК).

Литература

1. Алпатов Б.А., Бабаян П.В., Балашов О.Е., Степашкин А.И. Методы автоматического обнаружения и сопровождения объектов. Обработка изображений и управление. – М.: Радиотехника, 2008. – 176 с.
2. Визильтер Ю.В., Желтов С.Ю., Бондаренко А.В., Ососков М.В., Моржин А.В. Обработка и анализ изображений в задачах машинного зрения: курс лекций и практических занятий. – М.: Физматкнига, 2010. – 672 с.
3. Zitova B., Flusser J. Image registration methods: a survey //Image and Vision Computing 21–2003, p. 977–1000
4. Ташлинский А.Г. Оценивание параметров пространственных деформаций последовательностей изображений. – Ульяновск: УлГТУ, 2000. – 131 с.
5. Алпатов Б.А., Стротов В.В. Алгоритм оценки параметров геометрических преобразований изображений при мультиспектральном наблюдении // Цифровая обра-

ботка сигналов. – М.: РНТОРЭС им. А.С. Попова, 2010 – № 4. – С. 2-6.

6. Juan L., Gwun O. A comparison of sift, pca-sift and surf //International Journal of Image Processing (IJIP). – 2009. – Т. 3. – №. 4. – С. 143-152.

7. Bay H., Tuytelaars T., Van Gool L. Surf: Speeded up robust features //Computer Vision–ECCV 2006. – Springer Berlin Heidelberg, 2006. – С. 404-417.

8. Rublee E. et al. ORB: an efficient alternative to SIFT or SURF //Computer Vision (ICCV), 2011 IEEE International Conference on. – IEEE, 2011. – С. 2564-2571.

9. Lowe D. G. Distinctive image features from scale-invariant keypoints //International journal of computer vision. – 2004. – Т. 60. – №. 2. – С. 91-110.

10. Nistér D. Preemptive RANSAC for live structure and motion estimation //Machine Vision and Applications. – 2005. – Т. 16. – №. 5. – С. 321-329.

THE COMPARATION OF TWO STRUCTURAL IMAGE GEOMETRIC TRANSFORMATION PARAMETERS ESTIMATION ALGORITHMS

Strotov V.V.

In this article two structural image geometric transformation parameters estimation algorithms are considered. The first algorithm based on tracking a number of reference areas. The second is well-known algorithm based of image features (SIFT). The results of experimental examination and implementation guidance are given.