

# Оценка ориентации камеры на Манхэттенских сценах с использованием информации с оптического и инерционного датчиков

Е.В. Мясников<sup>1,2</sup>

<sup>1</sup>Самарский национальный исследовательский университет им. академика С.П. Королева, Московское шоссе 34А, Самара, Россия, 443086

<sup>2</sup>Институт систем обработки изображений РАН - филиал ФНИЦ «Кристаллография и фотоника» РАН, Молодогвардейская 151, Самара, Россия, 443001

**Аннотация.** В настоящей работе решение задачи оценки ориентации камеры выполняется при условии двух основных ограничений. В качестве первого ограничения выступает ограничение класса анализируемых сцен только Манхэттенскими сценами. В качестве второго - наличие акселерометра на мобильном устройстве. Для оценки характеристик предложенного решения был подготовлен набор данных, содержащий как сами фотографии, так и данные с акселерометра, а также информацию об истинной ориентации устройства. Экспериментальные исследования проведены с использованием собственного подготовленного набора данных.

## 1. Введение

Оценка ориентации камеры является одной из важнейших задач в трехмерном компьютерном зрении. Обычно оценка ориентации камеры производится с использованием калибровочных образцов, что требует участия человека. По этой причине особый интерес представляют автоматические методы определения ориентации.

Несмотря на присутствие в современных мобильных устройствах различных датчиков, таких как акселерометр, компас и т.п., их использование для оценки ориентации ограничено из-за невысокой точности и влияния шумов [1]. По этой причине, для определения ориентации камеры используется как оптическая информация, так и информация, получаемая с датчиков мобильных устройств.

В настоящей работе рассматривается метод определения ориентации камеры, основанный на анализе положения точек схода [2] - точек в плоскости перспективного изображения, в которых сходятся проекции взаимно параллельных линий трехмерного пространства. При этом задача решается при условии двух основных ограничений. В качестве первого ограничения выступает ограничение класса анализируемых сцен только Манхэттенскими сценами, в которых линии выровнены вдоль трех основных взаимно ортогональных направлений. Яркими примерами таких сцен являются фотографии городской застройки (указанными характеристиками могут обладать линии фасадов зданий), дорожных сцен (граница дорожного полотна, разметка, столбы), сцены внутри помещений (границы помещений, линии мебели, элементов декора - панели, плитки и т.д.). В качестве второго ограничения выступает наличие акселерометра на мобильном устройстве.

Определение ориентации камеры в настоящей работе производится последовательно в несколько этапов. На первом этапе с использованием информации, полученной с инерционного датчика, определяется направление силы тяжести, соответствующее направлению на первую точку схода. После этого положение первой точки схода уточняется по вертикальным линиям на оптическом изображении. На втором этапе определяются точки схода горизонтальных линий главных и боковых фасадов. По найденным с учетом данных инерционного датчика точкам схода определяется, собственно, ориентация камеры. Описываемый в работе метод определения точек схода основывается на идее, описанной в работе [3], согласно которой поиск горизонтальных точек схода можно выполнять вдоль линии горизонта, определяемой плоскостью, ортогональной направлению на вертикальную точку схода.

К сожалению, распространенные наборы данных для оценки методов поиска точек схода (см., например, [4]) не содержат информации с инерционных датчиков. По этой причине их использование для оценки методов, подобных описанному в настоящей работе, возможно только в режиме эмуляции показаний таких датчиков, как сделано, например, в [5, 6].

По указанной выше причине для оценки характеристик предложенного решения был подготовлен собственный набор данных, содержащий как сами фотографии, так и данные с акселерометра и информацию об истинной ориентации. Экспериментальные исследования проведены с использованием подготовленного набора данных.

Следует отметить, что первоначальная реализация алгоритма определения точек схода ранее описывалась в работах [5, 6]. Таким образом, в настоящей работе предложенный ранее подход развивается и исследуется, в том числе, с использованием подготовленного в рамках работы набора данных.

Работа организована следующим образом. В разделе 2 приводится описание разработанного метода определения ориентации камеры. В разделе 3 описывается методика моделирования и проводятся экспериментальные исследования. Работа завершается заключением и списком используемой литературы.

## 2. Метод

Как было сказано во введении, описываемый метод состоит в последовательном определении трех точек схода с последующим нахождением ориентации камеры. Общая схема метода представлена на Рис. 1.

Сначала выполняется предварительная обработка полученного с камеры изображения. В частности, осуществляется его масштабирование и поворот с точностью до 90 градусов в соответствии с информацией, полученной с инерционного датчика. При необходимости, полученный с датчика вектор трансформируется так, чтобы соответствовать направлению силы тяжести для повернутого изображения.

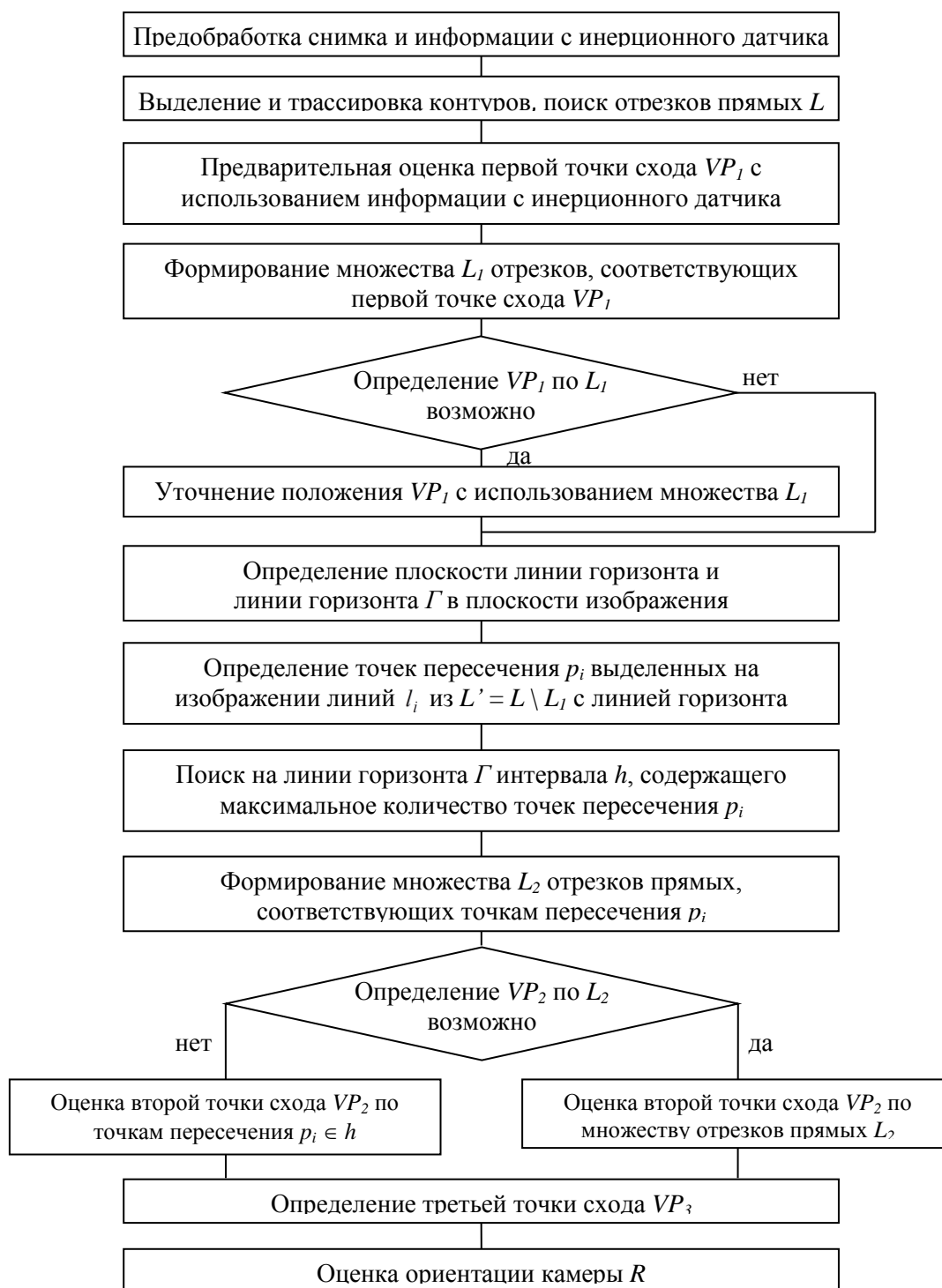
После предварительной обработки одним из известных методов, например, методом Canny, на изображении выделяются контура. Выполняется трассировка выделенных контуров и поиск прямых отрезков, входящих в их состав. Найденные отрезки формируют множество  $L$ , которое будет использоваться впоследствии для нахождения точек схода.

Далее информация, полученная с инерционного датчика, используется для предварительной оценки первой точки схода  $VP_1$ . При этом предполагается, что направление на первую точку схода соответствует с точностью до знака вектору силы тяжести.

По полученному направлению отбирается множество  $L_1$  таких отрезков, что соответствующие им прямые отклоняются от направления на  $VP_1$  не более, чем на заданный наперед угол.

В том случае, если отобранных отрезков оказывается достаточно, производится уточнение первой точки схода путем взвешенного суммирования точек, определяемых всевозможными отрезками из  $L_1$ . При этом отрезки большей длины имеют больший вес. Если определение  $VP_1$  по  $L_1$  невозможно, для дальнейшей обработки используется первоначальная оценка  $VP_1$ .

На следующем этапе направление на  $VP_1$  используется для определения плоскости линии горизонта, как плоскости, проходящей через начало координат моделируемой оптической системы и ортогональной направлению  $VP_1$  (уточненному вектору силы тяжести).



**Рисунок 1.** Общая схема метода.

Кроме плоскости, определяется и линия горизонта  $\Gamma$ , как проекция некоторой неортогональной к изображению линии в плоскости горизонта на плоскость изображения.

Далее для всех выделенных на изображении линий  $l_i \in L'$ , за исключением использованных ранее при определении точки схода  $VP_1$  ( $L' = L \setminus L_1$ ), определяются точки пересечения  $p_i$  с линией горизонта  $\Gamma$ . Производится поиск такого отрезка  $h$  (наперед заданного углового размера) на линии горизонта, в который попадает максимальное количество точек пересечения  $p_i$ .

После этого формируется множество  $L_2$  отрезков прямых, пересечения  $p_i$  которых с линией горизонта  $\Gamma$  попадают в указанный интервал  $h$ . В том случае, если отобранных отрезков оказывается достаточно, производится оценка второй точки схода путем взвешенного суммирования точек пересечения, определяемых всевозможными отрезками из  $L_2$ . Если определение  $VP_2$  по  $L_2$  невозможно, за положение  $VP_2$  берется взвешенная сумма точек пересечения  $p_i$  соответствующих линий с линией горизонта  $\Gamma$ : В обоих случаях отрезки большей длины имеют больший вес при определении  $VP_2$ .

После определения двух точек схода, третья находится, как вектор, ортогональный векторам, соответствующим первой и второй точке:  $V_3 = V_1 \times V_2$ .

После нахождения точек схода, ориентация камеры может быть найдена следующим образом:

$$R=[r_1 \ r_2 \ r_3],$$

где  $R$  - матрица вращения, векторы которой  $r_1, r_2, r_3$  рассчитываются как

$$r_1 = mK^{-1}VP_1, \ r_2 = mK^{-1}VP_2, \ r_3 = r_1 \times r_2,$$

где  $m$  - масштабный коэффициент,  $K$  – матрица внутренних параметров камеры, содержащая информацию о фокусном расстоянии, размерах пикселя, наклоне, сдвиге центра изображения относительно оптической оси.

В целом, предложенный метод представляет собой развитие описанного нами ранее метода [5, 6], основная идея которого [3] состоит в поиске горизонтальных точек схода вдоль линии горизонта, определяемой плоскостью, ортогональной направлению на вертикальную точку схода. По сравнению с предыдущей реализацией изменениям подверглись как отдельные этапы метода (поиск отрезков на контурах теперь выполняется по критерию максимального отклонения, взвешенное суммирование учитывает длины отрезков прямых, отстоящих друг от друга на достаточном расстоянии, уточнение второй точки схода выполняется без использования гистограмм), так и общая схема метода (теперь содержит ветвления, повышающие надежность определения точек схода, а также собственно этап оценки ориентации).

### 3. Эксперименты

Для исследования описанного выше метода, использовался собственный специально подготовленный набор данных. Указанный набор был собран с помощью смартфона Huawei Honor 9 lite. Его камера имеет сенсор CMOS BSI с диафрагмой  $f / 2.2$ , фокусным расстоянием 3.46 мм и выдает цветное изображение 12.98 МП. Для сбора изображений и данных инерционного датчика мы разработали Android приложение, в котором хранятся как полученные изображения, так и настраиваемое количество показаний акселерометра, зарегистрированных до момента снимка.

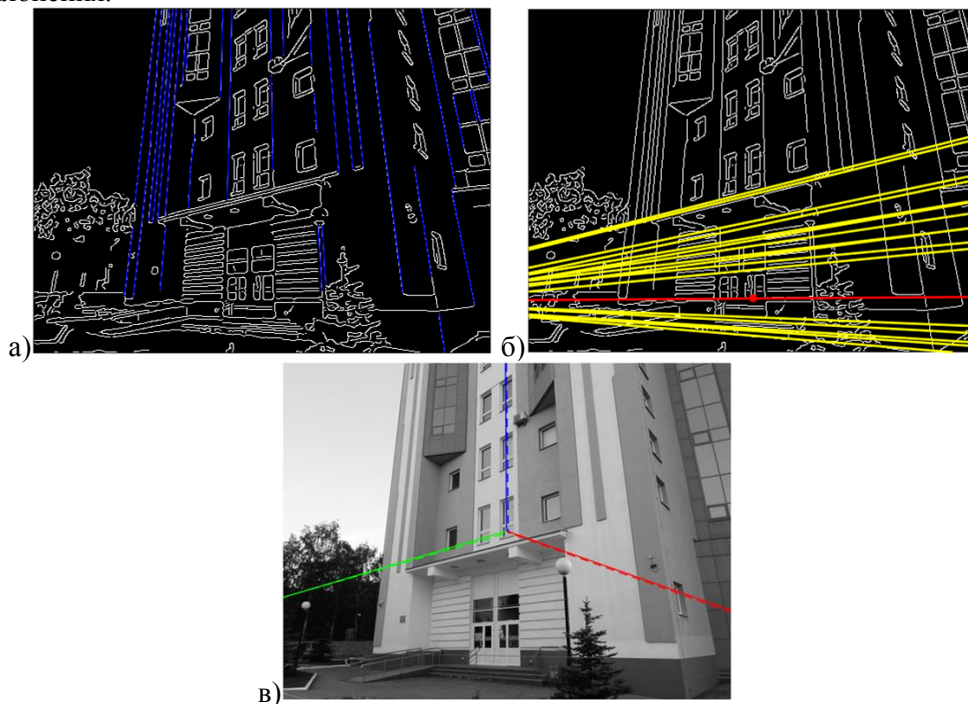
Для получения информации об истинном положении, производилось ручное выделение нескольких (от 3 до 7 для каждой точки схода) линий, которые надежно определяют направления к истинным точкам схода. Эта процедура была выполнена при 2-кратном увеличении, и в качестве истинных точек схода рассматривались нормализованные точки схода, полученные с использованием выделенных линий. На данный момент описанный набор данных состоит из 40 изображений зданий с соответствующими им данными инерционного датчика и данными об истинной ориентации.

Пример, демонстрирующий различные этапы работы предложенного метода, показан на рис. 2.

Для оценки качества разработанного метода моделирование выполнялось по следующей схеме:

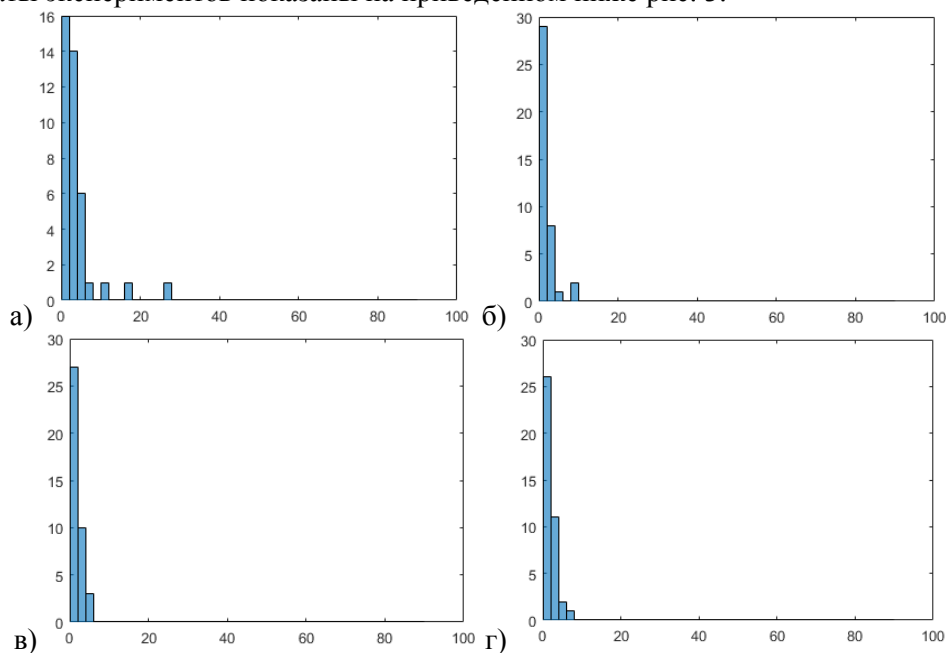
- для каждого изображения из сформированного набора данных выполнялось определение трех точек схода и ориентации камеры относительно изображенного на фотографии здания;
- с использованием информации об истинном положении точек схода для каждой точки схода рассчитывалась ошибка как угловое отклонение направления на оцененную точку схода от истинного направления;

- по полученным данным для каждой точки схода строилась гистограмма углового отклонения найденных точек от их истинных значений, а также вычислялось среднее значение такого отклонения.



**Рисунок 2.** Пример работы метода: а) выделенные контура (белым) и множество отрезков, соответствующих первой точке схода (синим); б) линия горизонта (красным) и множество прямых, определяющих вторую точку схода (красным); в) направления на истинные (пунктирными линиями) и оцененные (сплошными линиями) точки схода.

Результаты экспериментов показаны на приведенном ниже рис. 3.



**Рисунок 3.** Оценка качества работы метода. Гистограммы углов отклонений направлений на точки схода от их истинных значений (в градусах): а) первая точка схода, определенная по сведениям от инерционного датчика; б) первая точка схода, уточненная по оптическому изображению; в) вторая точка схода; г) третья точка схода.

Каждая из представленных на рисунке гистограмм показывает угловое отклонение положения оцененной точки схода от истинного положения. В идеальном случае подобная гистограмма должна иметь единственный столбец в левой части (первый), что означает минимальные отклонения точки схода от истинных значений для всех тестовых изображений. Как видно из приведенных рисунков в большинстве случаев положение трех точек схода выполнено с отклонением до  $2^\circ$ , при этом отклонение превысило  $4^\circ$  лишь для 3 изображений из 40. Средние значения ошибок при этом составили:  $1.69^\circ$ ,  $1.54^\circ$ , и  $1.88^\circ$  для первой, второй и третьей точек, соответственно.

Следует отметить, что использование только сведений от инерционного датчика (см. гистограмму на Рис. 1(a)) обеспечивало больший уровень ошибок при определении направления уже на первую точку схода. Среднее значение ошибки составило  $3.7^\circ$  при использовании только инерционного датчика против  $1.69^\circ$  при уточнении по оптическому изображению. Таким образом, точность работы алгоритма может быть повышена в условиях зашумления показаний вектора тяжести путем подбора параметров. Другим путем повышения точности может быть использование ранее полученных оценок при обработке видеопотока, что является предметом будущих исследований.

#### 4. Заключение

В работе предложен и исследован метод автоматического определения ориентации камеры на Манхэттенских сценах с использованием информации с оптического и инерционного датчиков. Для исследования разработанного метода был сформирован набор данных, содержащий цифровые снимки зданий, показания инерционных датчиков, а также сведения об истинном положении точек схода, полученные путем тщательной ручной разметки исходных изображений. Описанный метод прост в реализации, нетребователен к вычислительным ресурсам. Его использование позволяет более чем в 2 раза снизить средний уровень ошибок при определении ориентации по сравнению с инерционным датчиком.

В качестве направления дальнейших работ планируется расширить метод для оценки ориентации и положения камеры при работе с видеопотоком.

#### 5. Благодарности

Работа выполнена при финансовой поддержке РФФИ в рамках научного проекта № 17-29-03190-офи-м в частях «2. Методы» - «3. Эксперименты» и Министерства науки и высшего образования РФ в рамках госзадания ФНИЦ «Кристаллография и фотоника» РАН в частях «1. Введение» и «4. Заключение».

#### 6. Литература

- [1] Мясников, В.В. Исследование зависимости точности одновременной реконструкции сцены и позиционирования камеры от погрешностей, вносимых датчиками мобильного устройства / В. В. Мясников, Е. А. Дмитриев // Компьютерная оптика. – 2019. – 43(3). – С. 492-503.
- [2] Caprile, B. Using vanishing points for camera calibration / B. Caprile, V. Torre // International Journal of Computer Vision – 1990. – Vol. 4(2). – P. 127-139.
- [3] Angladon, V. The toulouse vanishing points dataset / V. Angladon, S. Gasparini, V. Charvillat // Proceedings of the 6th ACM Multimedia Systems Conference (MMSys) – Portland, United States, 2015.
- [4] Denis, P. Efficient Edge-Based Methods for Estimating Manhattan Frames in Urban Imagery / P. Denis, J.H. Elder, F. Estrada // Proc. European Conference on Computer Vision. – 2008. – Vol. 5303. – P. 197-211.
- [5] Мясников, Е.В. Автоматический поиск точек схода на мобильных устройствах / Е.В. Мясников // V Международная конференция и молодёжная школа «Информационные технологии и нанотехнологии» (ИТНТ) – Самара: Новая техника, 2019. – С. 451-457.
- [6] Myasnikov, E. Automatic search for vanishing points on mobile devices / E. Myasnikov // CEUR Workshop Proceedings. – 2019. – Vol. 2391. – P. 216-221.

# Assessment of camera orientation in Manhattan scenes using information from optical and inertial sensors

E.V. Myasnikov<sup>1,2</sup>

<sup>1</sup>Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

<sup>2</sup>Image Processing Systems Institute of RAS - Branch of the FSRC "Crystallography and Photonics" RAS, Molodogvardejskaya street 151, Samara, Russia, 443001

**Abstract.** In this paper, the assessment of the camera orientation is performed under two main limitations. The first one is the restriction of the class of analyzed scenes to Manhattan scenes only. The second limitation is the presence of an accelerometer on a mobile device. To assess the characteristics of the proposed solution, a dataset was prepared containing both the photographs and data from the accelerometer, as well as information about the true orientation of the device. Experimental studies were conducted using the prepared dataset.