

ГИРОСКОП НА ОСНОВЕ МОНОКУЛЯРНОЙ КАМЕРЫ

Представлен метод определения трехмерной ориентации монокулярной камеры с использованием точек схождения перспективы (ТСП), обнаруженных в последовательностях изображений. Устойчивое обнаружение ТСП в реальном времени осуществляется с использованием стандартного метода выделения сегментов линий (СЛ) и адаптивного алгоритма RANSAC. ТСП и соответствующие им направления, обнаруженные в последовательных кадрах, связываются друг с другом для построения последовательности кватернионов ориентации, которая обрабатывается расширенным фильтром Калмана. Экспериментами с мобильным телефоном показано, что точность предложенного метода сопоставима с точностями методов, использующих механические датчики движения потребительского класса точности.

1. Введение

При наличии недорогих миниатюрных видеодатчиков малого энергопотребления возрастает интерес к их использованию в системах позиционирования и навигационных приложениях с использованием таких цифровых устройств, как мобильные телефоны. Существует целый ряд работ в области компьютерного зрения, посвященных так называемой проблеме *собственного движения* (*egomotion*), предполагающей выделение последовательности изображений при трехмерном движении камеры (вращение с шестью степенями свободы и параллельное перемещение), установленной на роботе или транспортном средстве. Большинство этих методов основано на обнаружении базисных элементов изображения (точки, линии и так далее) и слежении за ними в потоке изображений. Такие методы, как одновременная локализация и картирование (SLAM) и определение структуры объекта в процессе движения (SfM), используют визуальные и другие данные для построения трехмерной модели окружающей среды, относительно которой описано движение камеры [10].

В этой работе мы сосредоточимся на разработке датчика изображения для определения трехмерной ориентации, то есть своеобразного гироскопа. Поскольку погрешность обычных гироскопов, основанных на инерциальных датчиках, накапливается со временем, навигационной системе необходимы не имеющие дрейфа источники данных, такие как датчики, основанные на анализе изображения [19, 22]. В принципе гироскоп на основе камеры не имеет дрейфа и таким образом должен служить хорошим дополнением к инерциальному гироскопу.

Наша система основана на отслеживании ТСП, которые были идентифицированы с помощью ЛС, обнаруженных на изображениях. Использование только ТСП делает слежение за ориентацией объекта значительно проще по следующим причинам:

Хуттунен Вилли. Студент Технического университета г. Тампере, Финляндия.

Пише Роберт. Доктор наук, профессор Технического университета г. Тампере.

Научный редактор перевода Б.С.Ривкин.

1. ТСП порождаются свойствами линейных объектов, которые различимы и в изобилии присутствуют в изображениях архитектурных элементов внутри помещений и в условиях плотной городской застройки, где определение местоположения с помощью глобальной навигационной спутниковой системы (GNSS) неудовлетворительно.
2. Число интересующих нас (ортогональных) ТСП ограничено, так как любой объект пространства в трехмерном мире может содержать самое большее три взаимно ортогональных направления.
3. ТСП связаны только с ориентацией камеры, а не с ее положением.
4. Методы на основе ТСП являются робастными относительно нестационарных объектов (людей, транспортных средств), перемещающихся в кадрах.

ТСП используются в таких приложениях, как навигация автономных транспортных средств [17], реконструкция архитектурных видов [1, 12, 20, 21] и коррекция фотографий [8]. Аналогично нашему подходу большинство исследований предполагает, что, хотя истинные характеристики камеры известны и зафиксированы, ТСП также могут использоваться для их калибровки [3, 11, 16, 18]. Гироскоп на основе видеокамеры, использующий ТСП в статье Кесслера и др. [14], подобен нашему, за исключением того, что в статье [14] не фильтруются оценки ориентации и определяются только линия горизонта и центральная ТСП.

Эта статья имеет следующую структуру. В разделе 2 мы определяем модель камеры как камеры-обскуры, вводим ТСП алгебраически и описываем их основные свойства, используя методы проективной геометрии. В разделе 3 описано совместное использование методов обнаружения СЛ и кластеризации линий на основе алгоритма RANSAC (RANdom Sample Consensus – метод оценки параметров модели на основе случайных выборок), применяемое для обнаружения ТСП. Метод получения достоверных параметров ориентации по обнаруженным ТСП и их фильтрации представлен в разделе 4. В разделе 5 приведены результаты эксперимента с использованием мобильного телефона, содержащего камеру и инерциальные датчики. Заключение и обсуждение дальнейших направлений исследований представлены в разделе 6.

2. Геометрия ТСП

Теория ТСП излагается обычно с использованием методов *проективной геометрии*, изучающих свойства геометрических объектов под воздействием проективных преобразований. Такое преобразование осуществляет и фото- или видеокамера, которая проецирует трехмерную сцену на двухмерную плоскость изображения. В проективной геометрии точки представлены в *однородных* координатах. Например, точка в двухмерном евклидовом пространстве $(x, y) \in \mathbb{R}^2$ может быть преобразована в однородную ей точку в проективном двухмерном пространстве $\mathbb{P}^2 \subseteq \mathbb{R}^3$ простым добавлением компоненты 1 в конец евклидова представления $(x, y, 1) \in \mathbb{P}^2$. В проективном пространстве \mathbb{P}^2 точки $(x, y, 1)$ и $(\alpha x, \alpha y, \alpha)$ эквивалентны для любого α , отличного от нуля. Важное подмножество \mathbb{P}^2 состоит из точек, имеющих вид $(x, y, 0)$. Эти точки известны, как идеальные точки, они соответствуют точкам, расположенным бесконечно далеко от начала координат. Подобным образом в проективном пространстве \mathbb{P}^3 идеальные точки имеют вид $(x, y, z, 0)$.

Для того чтобы определить ТСП алгебраически, мы должны определить

проективное преобразование, используемое для моделирования камеры. Одна из самых простых и обычно используемых моделей для конечной проективной камеры – модель булавочного отверстия (камеры-обскуры) [10]:

$$P: \mathbb{P}^3 \mapsto \mathbb{P}^2, \quad P(\mathbf{X}) = KR[I_3, -t] \mathbf{X}, \quad (1)$$

где $K \in \mathbb{R}^{3 \times 3}$ – матрица калибровки камеры, $R \in \mathbb{R}^{3 \times 3}$ – матрица ортогонального поворота и $t \in \mathbb{R}^3$ – вектор смещения камеры в мировой системе координат (рис. 1). Матрица калибровки K является верхней треугольной матрицей, задающей характерные параметры камеры:

$$K = \begin{bmatrix} \alpha f & s & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix},$$

где f – фокусное расстояние, p_x, p_y задают главную точку («центр» камеры), α – соотношение размеров сторон пиксела и s – коэффициент асимметрии камеры. Большинство цифровых камер на основе приборов с зарядовой связью (ПЗС) имеют квадратные пиксели ($\alpha = 1$), нулевую асимметрию ($s = 0$) и главную точку, расположенную близко к центру изображения. Также целесообразно предположить, что характерные параметры камеры не изменяются в течение времени.

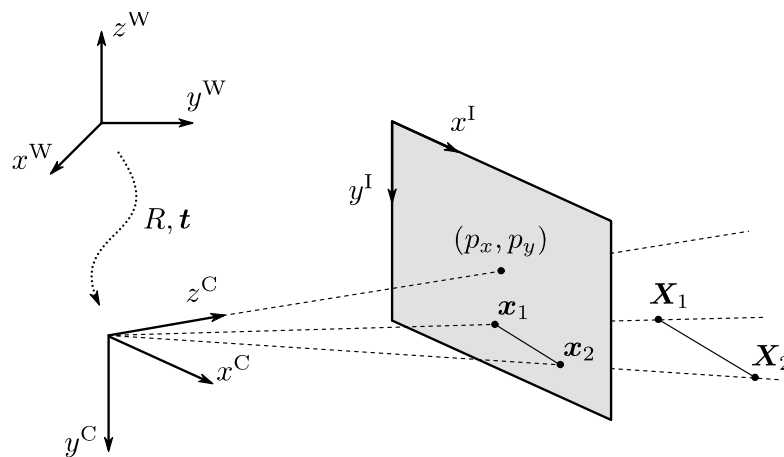


Рис. 1. Модель камеры-обскуры

Поворот камеры R и смещение t определяют преобразование координат из мировой системы W в систему координат C . Точка в системе координат камеры C может быть спроецирована на плоскость изображения I , если известна матрица калибровки камеры K . Сочетание этих двух преобразований приводит к модели камеры (1), которая точку $\mathbf{X} \in \mathbb{P}^3$ в мировой системе координат W проецирует в точку $\mathbf{x} \in \mathbb{P}^2$ на плоскости изображения I

Рассмотрим в трехмерном пространстве линию, которая проходит через точку $\mathbf{A} = (a_x, a_y, a_z, 1)^T$ с вектором направления $\mathbf{D} = (d^T, 0)^T, d \in \mathbb{R}^3$. Уравнение линии имеет вид

$$\mathbf{X}(\lambda) = \mathbf{A} + \lambda \mathbf{D}, \quad \lambda \in \mathbb{R}.$$

Когда эта линия спроецирована через камеру-обскуру P на плоскость изо-

бражения \mathbb{P}^2 , получим

$$\mathbf{x}(\lambda) = P(\mathbf{X}(\lambda)) = P(\mathbf{A}) + \lambda P(\mathbf{D}) = \mathbf{a} + \lambda K R \mathbf{d},$$

где $\mathbf{a} = P(\mathbf{A})$ – изображение линии \mathbf{A} . ТСП $\mathbf{v} \in \mathbb{P}^2$, соответствующая направлению \mathbf{d} , является предельной точкой спроецированной линии $\mathbf{x}(\lambda)$ при λ , стремящейся к бесконечности:

$$\mathbf{v} = \lim_{\lambda \rightarrow \infty} \mathbf{x}(\lambda) = \lim_{\lambda \rightarrow \infty} (\mathbf{a} + \lambda K R \mathbf{d}) = K R \mathbf{d}. \quad (2)$$

Когда ТСП (2) представлена в связанной с камерой системе координат ($R = I_3$, $\mathbf{t} = 0$), получаем

$$\mathbf{v} = K \mathbf{d}. \quad (3)$$

Из (2) можно заключить, что ТСП \mathbf{v} не зависит от положения камеры. Кроме того, уравнение (3) показывает, что есть взаимно однозначное соотношение между векторами направления в трехмерном пространстве и ТСП на плоскости изображения. Две ТСП \mathbf{v}_1 и \mathbf{v}_2 считаются ортогональными, если ортогональными являются соответствующие направления линий \mathbf{d}_1 и \mathbf{d}_2 . ТСП $\mathbf{v} \in \mathbb{P}^2$ (также идеальная точка), называется *бесконечной ТСП*. Аналогично неидеальные ТСП именуется *конечными ТСП*.

3. Обнаружение ТСП

Как отмечалось в предыдущем разделе, существует прямое соответствие между ориентацией камеры в мировой системе координат и ТСП, обнаруженными на плоскости изображения. Задача определения ориентации камеры поэтому может рассматриваться как задача обнаружения ТСП. Ориентация камеры имеет три степени свободы, следовательно, мы должны обнаружить три ТСП в каждом изображении. В подразделе 4.1 показано, что в предположении о взаимной ортогональности трех основных ТСП достаточно обнаружения только двух из этих точек, чтобы однозначно описать поворот камеры.

Обнаружение ТСП может рассматриваться как поиск таких точек в изображении, где проекции линий пересекаются. Наш метод обнаружения ТСП основан на обнаружении СЛ в изображениях и может быть разделен на две части:

1. *Обнаружение сегмента*. Детектор СЛ использует в качестве входного сигнала изображение в градациях серого и формирует список СЛ, обнаруженных в изображении.
2. *Кластеризация линий*. Обнаруженные СЛ объединяются в группы линий с общими точками пересечения. Затем отбираются группы с наибольшим числом СЛ и обнаруженные точки схождения назначаются точками пересечения внутри этих групп.

Следующие два подраздела описывают эти шаги более подробно.

3.1. Обнаружение СЛ

За несколько прошедших десятилетий были предложены многочисленные алгоритмы для обнаружения линий. Большинство методов основано на преобразовании Хо [6] или анализе связанных компонент ориентаций градиента изображения [2]. Мы используем метод Джиои и др. [9], который относится ко второй группе. Этот метод имеет линейную зависимость сложности от размера

изображения, но уступает по быстродействию методам, основанным на преобразовании Хо. С использованием определенного масштабирования изображения мы обнаружили, что метод Джиои может использоваться при работе в реальном времени. На рис. 2, б показан типичный выходной сигнал детектора линий [9] при исследованиях внутри помещения.

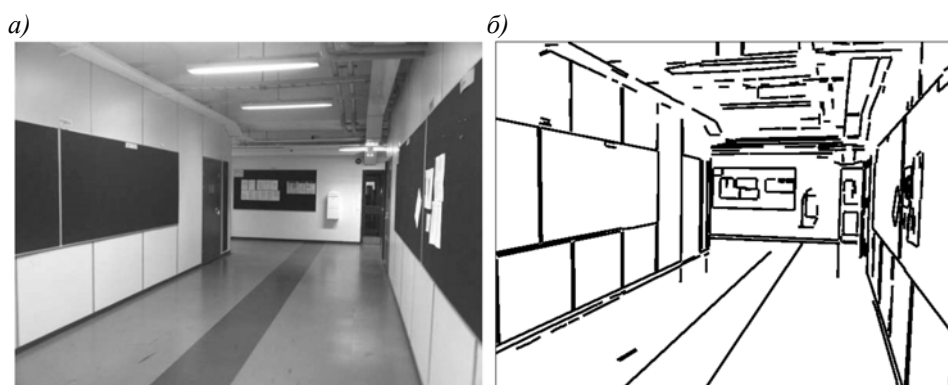
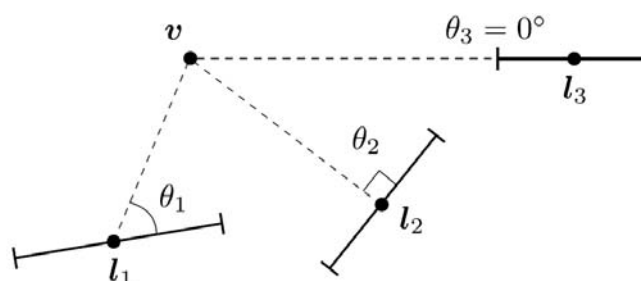


Рис. 2. Обнаружение СЛ

3.2. Кластеризация линий

После того как СЛ были выделены из изображения, используется алгоритм RANSAC [7] для группировки СЛ таким образом, чтобы их линии проходили через общую точку пересечения. Эти точки пересечения расцениваются как обнаруженные ТСП. Алгоритм RANSAC определяет одну группу линий за один прогон, поэтому, чтобы обнаружить три ТСП, алгоритм применяется три раза. Между каждым расчетом СЛ, соответствующие уже обнаруженным ТСП, удаляются. Мы не делаем различия между конечными и бесконечными ТСП во время кластеризации линий. В наших целях бесконечные ТСП достаточно хорошо аппроксимируются конечными, расположенными далеко от центра изображения.

Рис. 3. Углы между линейными сегментами и ТСП v

На каждом шаге алгоритма RANSAC пара СЛ выбирается случайным образом. Точка пересечения этой пары рассматривается как потенциальная ТСП. Чтобы оценить качество потенциальной ТСП, каждый из обнаруженных СЛ маркируется как «внутренний» или «внешний» на основании того, как хорошо он соответствует данной ТСП. Функция расстояния $d(v, l)$ между ТСП v и СЛ l определяется как угол между СЛ l и линией, проходящей через v и середину l (рис. 3). СЛ l_i считается «внутренним», если его расстояние от ТСП $d(v, l_i) < \theta_i$, для некоторого заданного порогового уровня θ_i .

Число итераций алгоритма RANSAC определяется адаптивно следующим образом. Обозначим r как отношение внутренних и внешних СЛ, соответствующих ТСП \mathbf{v} в наборе обнаруженных СЛ. Определим число итераций k , необходимых, чтобы гарантировать с вероятностью p , что по крайней мере одна из пар линии, случайно выбранных алгоритмом RANSAC, не содержит внешних СЛ. Вероятность выбора k пар СЛ по крайней мере с одним внешним в каждой из них составляет $(1 - r^2)^{k \cdot k}$. Число итераций, необходимых, чтобы гарантировать вероятность ниже порога $1 - p$:

$$k \geq \frac{\log(1-p)}{\log(1-r^2)}. \quad (4)$$

Истинное значение r неизвестно, но оно может быть аппроксимировано снизу числом СЛ в самой большой группе, найденной на более ранних итерациях, поделенным на число всех найденных СЛ.

RANSAC вырабатывает грубые оценки ТСП и соответствующих им наборов внутренних СЛ. В следующем подразделе описано, как дополнительная информация, содержащаяся в наборах внутренних СЛ может использоваться для дальнейшего уточнения оценок ТСП.

3.3. Уточнение ТСП

Рассматривая набор внутренних СЛ $\{l_1, \dots, l_n\}$, $l_i \in \mathbb{R}^3$, соответствующий ТСП \mathbf{v} , получаем следующее уравнение:

$$[l_1, \dots, l_n]^T \mathbf{v} = \mathbf{0}. \quad (5)$$

Переопределенную систему (5) можно решить для нетривиального \mathbf{v} (базиса нулевого пространства $[l_1, \dots, l_n]^T$), используя сингулярное разложение. Вместо того чтобы непосредственно решать систему (5), Чиполла и Бойер [4] предлагают сначала преобразовать линии l_i в нормализованные координаты изображения

$$L_i = \frac{K^{-1}l_i}{\|K^{-1}l_i\|}, \quad i = 1, \dots, n,$$

где K – известная матрица калибровки камеры. Нормализованные координаты изображения соответствуют камере с фокусным расстоянием 1 и главной точкой, расположенной в начале координат изображения. Вместо того чтобы решать систему (5) для ТСП \mathbf{v} , мы решаем для соответствующего стремящегося к нулю направления \mathbf{d} следующую систему:

$$[L_1, \dots, L_n]^T \mathbf{d} = L\mathbf{d} = \mathbf{0}. \quad (6)$$

Нулевые направления могут быть преобразованы обратно в ТСП путем умножения их слева на матрицу калибровки камеры K . Пример групп линий, соответствующих уточненным оценкам ТСП, найденных в кадре, снятом внутри помещения, показан на рис. 4.

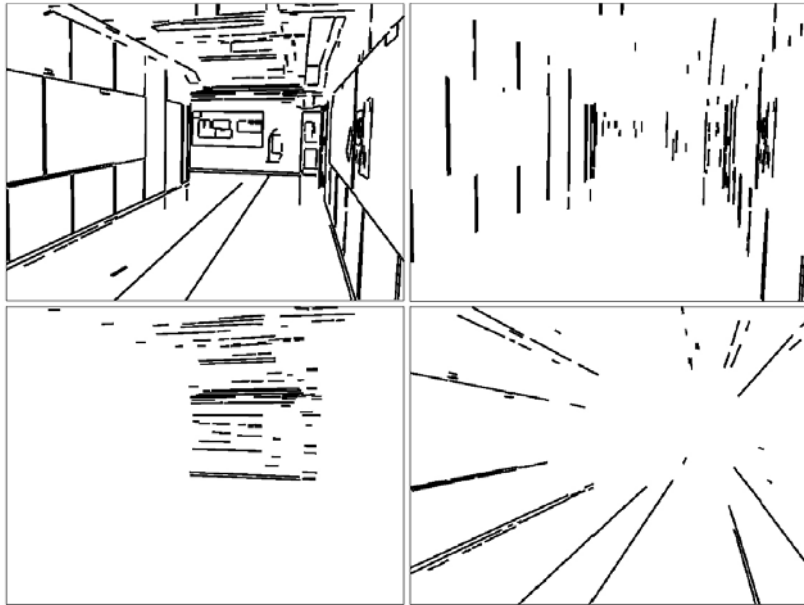


Рис. 4. Результаты кластеризации линий на основе алгоритма RANSAC

На верхнем левом изображении видны все обнаруженные СЛ. На остальных трех изображениях видны группы линий, соответствующие трем главным ортогональным ТСП, обнаруженным в кадре. Верхнее правое и нижнее левое изображения соответствуют вертикальным и горизонтальным ТСП, расположенным далеко от центра изображения. Нижнее правое изображение соответствует ТСП в конце коридора.

4. Слежение за ориентацией

В этом разделе представлен простой метод вычисления ориентации камеры по данным о двух или трех взаимно ортогональных ТСП. Также описан фильтр Калмана для оценки реального поворота камеры по последовательности определения ориентации на основе теории кватернионов.

4.1. Определение ориентации по ТСП

Уравнение (2) связывает ТСП с ориентацией камеры и фактическими направлениями линий, найденных в кадре. Предположим, что в кадре существуют три главных направления линий, которые являются взаимно ортогональными, и что обнаруженные ТСП соответствуют этим направлениям. Это предположение разумно для кадров внутри прямоугольных помещений. Обнаружение ТСП с помощью алгоритма RANSAC, описанное в предыдущих разделах, приводит к нахождению трех потенциальных ТСП независимо от того, сколько направлений ортогональных линий можно увидеть в изображении (например, изображение стены может иметь самое большее две ортогональные ТСП). При предположении ортогональности эти ТСП должны быть ортогональными и соответствующие нулевые направления

$$d_i = \frac{K^{-1}v_i}{\|K^{-1}v_i\|}, \quad i = 1, 2, 3$$

должны формировать ортонормированный базис в \mathbb{R}^3 . Если только два из нулевых направлений являются приблизительно ортогональными (одна из обнаруженных ТСП ложная), третье направление можно определить на их основании, используя векторное произведение.

Из-за приближительного характера обнаружения ТСП матрица $D = [d_1, d_2, d_3]$ не обязательно является ортогональной. Можно показать [13, с. 431-432], что ближайшая (в смысле нормы Фробениуса) к D ортогональная матрица задается

$$\hat{D} = UV^T,$$

где $D = U\Sigma V^T$ является разложением по сингулярным числам матрицы D .

При двух заданных ортонормированных базисах $D_1 = [d_1, d_2, d_3]$ и $D_2 = [d'_1, d'_2, d'_3]$ их относительная матрица поворота определяется как

$$R = D_2 D_1^T.$$

Для любого изображения матрица нулевых направлений D может быть выбрана разными способами. Знаки и порядок обнаруженных нулевых направлений d_i произвольны и могут меняться в зависимости от системы координат. Чтобы избежать неоднозначности в матрицах нулевых направлений D_i , мы переставляем и изменяем знаки их столбцов таким способом, чтобы они как можно больше были похожи на матрицу нулевых направлений, найденную в предыдущей системе координат. Поясним сказанное подробнее. Предположим, что матрица $D_1 = [d_1, d_2, d_3]$ состоит из нулевых направлений, найденных в предыдущей системе координат $i-1$, а матрица $D_2 = [d'_1, d'_2, d'_3]$ составлена из нулевых направлений, найденных в текущей системе координат i . Чтобы устранить неоднозначность в D_2 и гарантировать, что это истинная матрица поворота, мы связываем ее с матрицей D_1 . Чтобы осуществить это, мы вычисляем матрицу S следующим образом:

$$S = D_2^T D_1 = \begin{bmatrix} d_1^T d'_1 & d_1^T d'_2 & d_1^T d'_3 \\ d_2^T d'_1 & d_2^T d'_2 & d_2^T d'_3 \\ d_3^T d'_1 & d_3^T d'_2 & d_3^T d'_3 \end{bmatrix}.$$

Мы хотим, чтобы порядок столбцов матрицы D_2 соответствовал порядку столбцов матрицы D_1 . Чтобы определить соответствующий порядок, мы ищем максимальные абсолютные значения в строках $i = 1, 2$ матрицы S . Для каждой строки, если максимальное абсолютное значение строки i найдено в столбце j , меняют местами столбцы i и j в матрице D_2 , если же $i = j$, ничего делать не нужно. Кроме того, если знак S_{ij} отрицательный, умножаем столбец i перегруппированной матрицы D_2 на -1 . Изменение порядка и знаков столбцов D_2 не изменит того, что это ортогональная матрица. Если матрица D , найденная в первой системе координат, была связана с единичной матрицей I_3 , известно, что её определитель $\det D = +1$. Описанный процесс гарантирует, что это свойство сохраняется в последующих системах координат.

4.2. Фильтрация ориентации

В предыдущих разделах описано, как получить прямую, содержащую шумовую составляющую, оценку ориентации камеры относительно изучаемого объ-

екта. Чтобы улучшить оценку ориентации, используем расширенный фильтр Калмана. Помимо сглаживания шума, фильтрация позволяет продлить процесс оценивания ориентации камеры в тех случаях, когда не удается обнаружить две ортогональные ТСП.

Представим ориентацию камеры как четырехмерный кватернион ориентации. Кватернионы ориентации легко могут быть преобразованы в матрицы поворота и углы Эйлера [5], но при этом являются более компактными, чем матрицы поворота, и не имеют особых точек, которые могут быть присущи углам Эйлера.

Для движущейся камеры необходимо моделировать параметры движения более высокого порядка, такие как скорость и ускорение. Включим оценку угловой скорости камеры в вектор состояния и будем моделировать угловые ускорения, как случайный шум. Полный вектор состояния для движущейся камеры x состоит из кватерниона ориентации камеры $q \in \mathbb{R}^4$ и вектора угловой скорости $\omega \in \mathbb{R}^3$

$$x = \begin{bmatrix} q \\ \omega \end{bmatrix}.$$

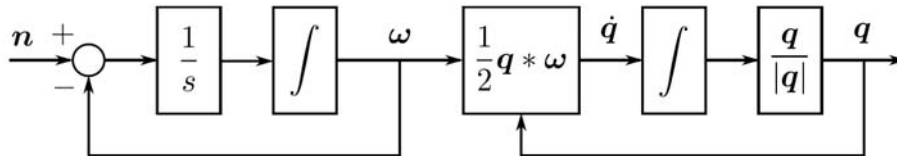


Рис. 5. Модель процесса фильтрации ориентации с использованием расширенного фильтра Калмана

Для последовательной фильтрации в процессе движения мы представляем непрерывное движение камеры в виде последовательности дискретных состояний, которые маркируем, используя нижний индекс k . Предполагаем, что временной шаг Δt между любыми двумя последовательными состояниями камеры является постоянным и равным обратному значению частоты кадров камеры. Допустим, что движение камеры описывается с помощью постоянной угловой скорости с неизвестными случайными гауссовыми угловыми ускорениями α_k с нулевым средним, возникающими на каждом временном шаге k

$$f(x_k, n) = \begin{bmatrix} q_k * q((\omega_k + n)\Delta t) \\ \omega_k + n \end{bmatrix},$$

где $n = \alpha\Delta t \sim N(0, Q)$ – импульс угловой скорости, который используется для моделирования неопределенности движения камеры, $q(\cdot)$ – преобразование поворотов вокруг осей (трехмерный вектор) к кватерниону ориентации. Модель измерения h отражает кватернионную часть вектора состояния плюс шум,

$$h(x_k) = Hx_k + r_k = q_k + r_k,$$

где $r_k \sim N(0, R_k)$ и $H = [I_3, 0_{4 \times 3}]$.

Стандартный алгоритм расширенного фильтра Калмана состоит из шагов прогнозирования и обновления. Шаг прогнозирования выглядит так:

$$\begin{aligned}x_k &= f(x_{k-1}, 0) \\P_k &= F_{k-1}P_{k-1}F_{k-1}^T + G_{k-1}QG_{k-1}^T,\end{aligned}$$

где

$$\begin{aligned}F_{k-1} &= \left. \frac{\partial f}{\partial x} \right|_{x=x_{k-1}, n=0} \\G_{k-1} &= \left. \frac{\partial f}{\partial n} \right|_{x=x_{k-1}, n=0}\end{aligned}$$

Эти матрицы имеют достаточно громоздкий вид и потому не представлены здесь.

Шаг обновления записывается в виде:

$$\begin{aligned}z_k &= y_k - h(x_k) \\S_k &= H_k P_k H_k^T + R_k \\K_k &= P_k H_k^T S_k^{-1} \\x_k &= x_k + K_k z_k \\P_k &= (I - K_k H_k) P_k\end{aligned}$$

Когда при обнаружении ТСП не удастся выработать приемлемую ориентацию, мы пропускаем шаг обновления и используем только шаг прогнозирования.

5. Результаты эксперимента

Эксперименты проводились с использованием мобильного телефона Nokia N900. Эта модель телефона, как и многие другие современные смартфоны, имеет камеру с разрешением 5 мегапикселей для съемки фотографий или видео (720 точек при 25 кадрах в секунду). У этой модели телефона есть также внутренний инерциальный измерительный модуль (ИИМ) с трехосным акселерометром, трехстепенным гироскопом и магнитометром.

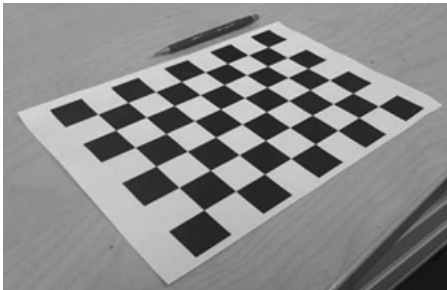


Рис. 6. Образец для калибровки

Матрица калибровки K для камеры была определена с использованием метода, подобного описанному в работе Zhang [23]. Процесс калибровки включал многократное фотографирование плоского об-

разца, показанного на рис. 6.

В наших экспериментах телефон использовался для видеосъемки с рук сцен внутри помещения, в то время как непосредственные определения параметров ориентации с помощью ИИМ регистрировались в фоновом режиме. Впоследствии видео было разделено на 680 отдельных кадров. СЛ и ТСП были обнаружены в каждом кадре автономно с использованием настольного компьютера, и предложенный в настоящей статье метод использовался для слежения за ориентацией камеры по последовательности изображений.

Видеокадры, снятые телефоном, имели исходное разрешение 848×480 пикселей, но, чтобы ускорить обнаружение СЛ, разрешение всех кадров перед обработкой было уменьшено в 2 раза. Код программы обнаружения СЛ написан на языке Си, кластеризация линий на основе алгоритма RANSAC, определение и фильтрация параметров ориентации выполнены, как неоптимизированный код MATLAB. Ясно, что обнаружение СЛ с использованием детектора [9] было «узким местом» с точки зрения быстродействия. Используя двухъядерный настольный компьютер с частотой 2.27 ГГц, мы смогли обрабатывать видео со средней скоростью 20 кадров в секунду, которая очень близка к собственной частоте кадров камеры (25 кадров в секунду). Получающиеся оценки параметров ориентации были подобны оценкам, полученным с помощью ИИМ (рис. 7).

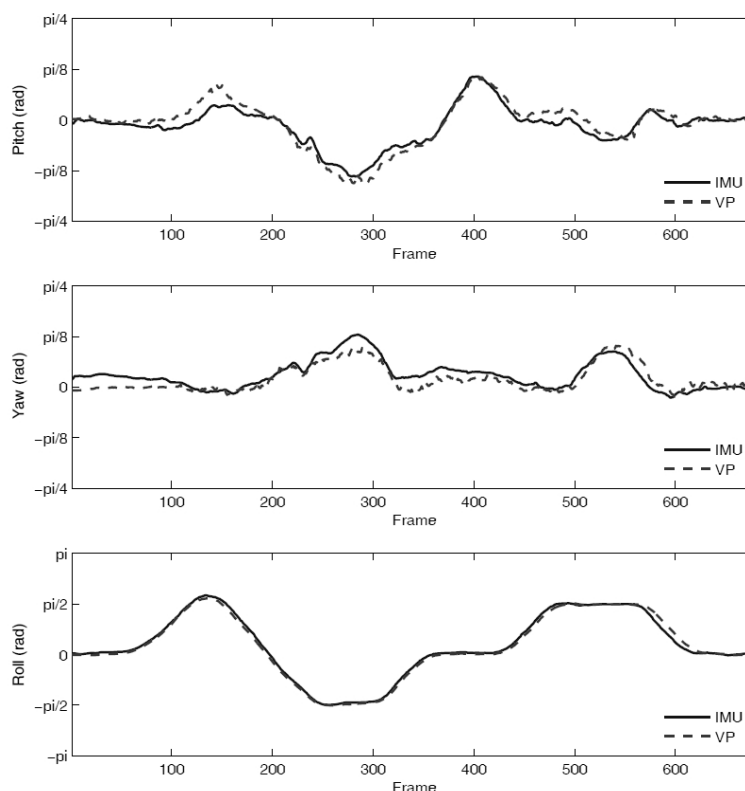


Рис. 7. Результаты эксперимента с карманной камерой.

Пунктирная линия представляет углы поворота, измеренные с использованием предложенного нами метода, сплошная линия – выходной сигнал встроенного ИИМ.

Pitch = килевая качка, Yaw = рыскание, Roll = бортовая качка, Frame = кадр

6. Выводы

В статье представлена концепция построения трехстепенного гироскопа, который работает, следя за ТСП, обнаруженными в отдельных кадрах последовательности изображений, снятых монокулярной камерой. Эксперименты показали, что система может следить за ориентацией камеры в режиме реального времени с точностью, сопоставимой с точностью ИИМ потребительского класса. Поскольку видеосистема в принципе не имеет дрейфа, у нее есть хороший потенциал для использования в качестве компонента интегрированной навигационной системы.

Самое слабое звено в представленной гироскопической системе на основе камеры – это обнаружение линий. Во-первых, изображения должны содержать достаточно характерных линий, чтобы определять ТСП. Таким образом, можно ожидать, что система будет хорошо работать главным образом в хорошо освещенной искусственно созданной окружающей среде, такой как внутренние помещения и городские условия. Линии трудно обнаружить в изображениях, которые размыты быстрым движением камеры. В этом случае гироскоп на основе камеры Кляйна и Драммонда [15], базирующийся на анализе вызванной движением размытости изображений, мог бы быть хорошей альтернативой или дополнением к предложенному решению. Наконец, поскольку обнаружение линии это часть нашей системы, требующая наибольших вычислительных затрат, самой важной областью дальнейших исследований является разработка более быстрых алгоритмов обнаружения ТСП по потоку изображений.

Благодарности

Эта работа была частично профинансирована корпорацией Nokia.

ЛИТЕРАТУРА

1. **M. Antone and S. Teller.** Automatic recovery of relative camera rotations for urban scenes. In Proceedings of IEEE Conference of Computer Vision and Pattern Recognition (CVPR'00), pages 282-289, Hilton Head, SC, USA, 2000.
2. **J. Burns, A. Hanson, and E. Riseman.** Extracting straight lines. IEEE Transactions on Pattern Analysis and Machine Intelligence, 8(4): 425-455, 1986.
3. **B. Caprile and V. Torre.** Using vanishing points for camera calibration. International Journal of Computer Vision, 4:127-140, 1990.
4. **R. Cipolla and E. Boyer.** 3D model acquisition from uncalibrated images. In Proceedings of the 1998 IAPR Workshop on Machine Vision Applications (MVA'98), pages 559-568, Chiba, Japan, 1998.
5. **J. Diebel.** Representing attitude: Euler angles, unit quaternions, and rotation vectors. Technical report, Stanford University, Palo Alto, CA, 2006.
6. **R. Duda and P. Hart.** Use of the Hough transformation to detect lines and curves in pictures. Communications of the ACM, 15:11-15, 1972.
7. **M. Fischler and R. Bolles.** Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM, 24(6):381-395, 1981.
8. **A. Gallagher.** Using vanishing points to correct camera rotation in images. In Proceedings of the Second Canadian Conference on Computer and Robot Vision (CRV'05), pages 460-467, Victoria, Canada, 2005.
9. **R. Gioi, J. Jakubowicz, J. Morel, and G. Randall.** LSD: A fast line segment detector with a false detection control. IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(4):722-732, 2010.
10. **R. Hartley and A. Zisserman.** Multiple View Geometry in Computer Vision. Cambridge University Press, New York, second edition, 2003.
11. **B. He and Y. Li.** Camera calibration from vanishing points in a vision system. Optics and Laser Technology, 40(3):555-561, 2008.
12. **F. Heuvel.** Vanishing point detection for architectural photogrammetry. International Archives of Photogrammetry and Remote Sensing, 32(5):652-659, 1998.
13. **R. Horn and C. Johnson.** Matrix Analysis. Cambridge University Press, 1985.
14. **C. Kessler, C. Ascher, N. Frietsch, M. Weinmann, and G. Trommer.** Vision-based attitude estimation for indoor navigation using vanishing points and lines. IEEE ION/PLANS 2010, pages 310-318, 2010.
15. **G. Klein and T. Drummond.** A single-frame visual gyroscope. In Proceedings of British Machine Vision Conference (BMVC'05), volume 2, pages 529-538, Oxford, 2005.
16. **J. Kořecká and W. Zhang.** Video compass. In Proceedings of European Conference on Computer Vision (ECCV'02), pages 657-673, 2002.

17. **Y. Lee, C. Nam, K. Lee, Y. Li, S. Yeon, and N. Doh. VPass.** Algorithmic compass using vanishing points in indoor environments. In Proceedings of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'09), pages 936-941, St. Louis, MO, USA, 2009.
18. **J. Lobo and J. Dias.** Vision and inertial sensor cooperation using gravity as a vertical reference. IEEE Transactions on Pattern Analysis and Machine Intelligence, 25(12):1597-1608, 2003.
19. **F. Mirzae.** A Kalman filter-based algorithm for IMU-camera calibration: Observability analysis and performance evaluation. IEEE Transactions on Robotics, 24(5), 2005.
20. **J. Montiel and A. Zisserman.** Automated architectural acquisition from a camera undergoing a planar motion. In Proceedings of International Symposium on Virtual and Augmented Architecture (VAA'01), pages 207-218, Dublin, Ireland, 2001.
21. **C. Rother.** A new approach to vanishing point detection in architectural environments. Image and Vision Computing, 20(9-10):647-655, 2002.
22. **S. Roumeliotis, A. Johnson, and J. Montgomery.** Augmenting inertial navigation with image-based motion estimation. In Proceedings of IEEE International Conference on Robotics and Automation (ICRA'02), pages 4326-4333, Washington D.C., USA, 2002.
23. **Z. Zhang.** Flexible camera calibration by viewing a plane from unknown orientations. In Proceedings of International Conference on Computer Vision (ICCV'99), pages 666-673, Corfu, Greece, 1999.

Abstract. We present a method for tracking the 3-axis orientation of a monocular camera using orthogonal vanishing points detected in individual frames of a sequence of images. Robust and real-time vanishing point detection is done using a standard line segment detection method and an adaptive RANSAC algorithm. Vanishing points and corresponding vanishing directions found in consecutive frames are associated with each other to produce a sequence of orientation quaternions, which is processed by an extended Kalman filter. Experiments with a consumer-level, handheld mobile device indicate that the accuracy of the proposed method is comparable with those of consumer-grade mechanical motion sensors.

Key words: Computer vision, indoor navigation, vanishing points, orientation estimation

Материал поступил 15.08.11