Аналитический подход к применению схемы RANSAC в задачах визуальной навигации в динамическом окружении

Книжов Кирилл Игоревич

1 Введение

Устойчивость к динамическому окружению - одна из ключевых задач в разработке визуальных систем навигации и VisualSLAM систем. Эта задача имеет множество подходов, однако наиболее распространенным является подход с применением итерационной схемы RANSAC (она вкратце описана в п.3). Эта схема предполагает вычисление навигационных параметров (ориентации и перемещения) на каждой итерации, что не является необходимым для решения задачи об исключении динамических объектов. В связи с этим в этой статье была предложена модификация схемы RANSAC, исключающая вычисление навигационных параметров на каждой итерации (п. 4).

2 Визуальная навигация

В данной работе будет рассматриваться один из основных подходов к визуальной навигации в своей вариации для стереопары видеокамер, который подразумевает следующий алгоритм:

1) Поиск особых точек (feature detecting) на "левом" изображении

2) Поиск этих точек (feature matching) на правом изображении

3) Вычисление пространственных координат этих точек

4) Отслеживание этих точек на видеопотоке (feature tracking)

5) Вычисление пространственных координат этих точек в

6) Вычисление изменения позиции видеокамеры по изменению пространственных координат точек(PnP problem)

Данный алгоритм является классическим и был множество раз описан в различных работах (см [1],[2]), также нередко можно встретить его вариацию и для монокамеры (см. [2],[3]), а потому мы не будем расписывать каждый его шаг. Остановимся чуть более подробно лишь на последнем пункте, который имеет непосредственное отношение к теме данной работы.

Задача PnP - задача определения позиции набора ключевых точек в текущий момент времени по их изображению, при известных координатах этих точек в начальный момент времени. Пусть ${}^{G}p_i$ и $d_i * {}^{G}b_i$ - координаты i-ой ключевой точки в начальный и текущий момент времени соответственно (в системе координат камеры), где b_i - измеренные по изображению единичные векторы, d_i - скалярные величины, ${}^{C}p_C \in \mathbb{R}^3$ - вектор параллельного переноса камеры, ${}^{G}_{C}C \in SO(3)$ - матрица вращения. Тогда задача PnP сводится к решению системы

$$ {}^{G}p_i = {}^{C}p_C + d_i {}^{G}_{C}C {}^{G}b_i, i = 1..n, \tag{1}$$

относительно неизвестных ${}^{G}_{C}C, d_i, {}^{C}p_C$. Стоит отметить, при $n = 3$ и задача полностью классифицирована и может быть решена аналитически (см. [4],[5]), а при $n \geqslant 4$ используются приближенные алгоритмы, такие как метод наименьших квадратов и различные его оптимизации (см., например, [6]).

Если нам известны изображения ключевых точек со стереопары видеокамер, мы можем считать, что нам известны коэффициенты d_i. Таким образом, система (1) принимает вид

$$ {}^{G}p_i = {}^{G}p_C + {}^{G}_{C}C {}^{C}p_i, i = 1..n, \tag{2}$$

где ${}^C p_i$ - измерения координат особых точек в текущий момент времени в системе координат камеры. Легко видеть, что данная система однозначно разрешима в случае $n = 3$ и неколлинеарности точек ${}^G p_i$ (при условии, что ${}^C p_i$ вообще могут быть получены из ${}^G p_i$ с помощью аффинного преобразования).

3 Визуальная навигация в динамическом окружении

В реальных приложениях окружение зачастую оказывается динамическим, что сильно уменьшает точность описанных выше алгоритмов. Таким образом возникает задача повышения устойчивости к так называемым "выбросам"(outliers). Классическим решением данной проблемы является применение итерационной схемы RANSAC (см. [7]). В общих чертах данная схема выглядит следующим образом:

1) случайным образом выбирается n особых точек изображения;

2) для этих n точек решается задача PnP;

3) по некоторому критерию полученное решение проверяется для всех точек;

4) если количество точек, которое удовлетворяет полученному решению, больше, чем у предыдущего лучшего по этому параметру, то данное решение выбирается текущим лучшим решением.

Количество итераций k в данном процессе определяется вероятностью получения корректного решения p и предполагаемого количества "инлаеров" по отношению к количеству всех точек w по формуле:

$$k = log(1-p)/log(1-w^n)$$

Данный подход широко используется в различных реализациях визуальных навигационных алгоритмов (программы VisualSLAM) и показывает хорошие результаты (см., например, [2], [3]), однако сам по себе не раскрывает особенностей конкретной задачи, а потому представляет естественный интерес вопрос об аналитических свойствах проблемы динамического окружения в задаче визуальной навигации.

Пусть $X = \{x_i\}_{i=1}^n$ - некоторое множество наблюдаемых точек пространства, для которых известны их координаты ${}^1 x_i$ и ${}^2 x_i$ в некоторой ортонормированной системе координат в момент времени t_1 и t_2 соответственно. Будем говорить, что множество $Y = \{y_i, i = 1..m\} \subseteq X$

подчинено одному аффинному преобразованию, если:

$$\exists C \in SO(3), T \in \mathbb{R}^3 : {}^2y_i = C\,{}^1y_i + T \;\forall i = 1..m,$$

где ${}^1y_i$ и ${}^2y_i$ - координаты точки y_i в момент времени t_1 и t_2 соответственно, а $SO(3)$ - специальная ортогональная группа вещественных матриц 3×3, т.е. множество таких матриц F, что $detF = 1$ и $\|Fx\| = \|x\| \;\forall x \in \mathbb{R}^3$ (в дальнейшем мы будем отождествлять эти матрицы с соответствующими им линейными операторами).

Тогда задачу исключения "выбросов"можно представить как задачу нахождения максимального по включению множества Y, подчиненного одному аффинному преобразованию (эти точки обычно называют "инлаерами"(inliers)). Для решения этой задачи нами был сформулирован следующий критерий подчиненности точек одному аффинному преобразованию.

4 Модификация схемы RANSAC с применением аналитического подхода

Схема RANSAC, описанная в п.3, подразумевает в шаге 2 вычисление аффинного преобразования на каждой итерации. Однако это не является необходимым для нахождения множества "инлаеров в связи с чем возникает естественное желание модифицировать схему RANSAC так, чтобы исключить из нее это вычисление. С этим нам поможет следующий критерий.

Утверждение 1. *Пусть X -множество точек, наблюдаемых в моменты времени t_1 и t_2, $y_i \in X, i = 1..3$ - множество точек, подчиненное одному аффинному преобразованию,не лежащих на одной прямой (т.е. $\exists C \in SO(3), T \in \mathbb{R}^3 \; : {}^2y_i = C\,{}^1y_i + T \;\forall i = 1..3$).Пусть точка $x \in X, x \neq y_i, i = 1..3$ такова, что*

$$\|{}^1a_i\| = \|{}^2a_i\| \;\forall i = 1..3 \tag{3}$$

и

$$sign(detA_1) = sign(detA_2), \tag{4}$$

где ${}^ka_i = {}^kx - {}^ky_i,\; k = 1,2,\; A_k = ({}^ka_1, {}^ka_2, {}^ka_3),\; k = 1,2.$ Тогда множество $\{x, y_1, y_2, y_3\}$ подчинено одному аффинному преобразованию.

Для доказательства данного утверждения нам потребуется две леммы.

Лемма 1. *Пусть* $C, \tilde{C} \in SO(3)$, $P \in \mathbb{R}^3$ *таковы, что*

$$Cy_i = \tilde{C}y_i + P, i = 1..3,$$

где $y_i \in \mathbb{R}^3$ 0 - *три различные точки, не лежащие на одной прямой. Тогда* $T = 0$ *и* $C = \tilde{C}$.

Доказательство:

Заметим, что в условиях леммы возможно 2 случая:

1) векторы y_i не лежат в одной плоскости;

2) векторы y_i лежат в одной плоскости.

В первом случае утверждение леммы очевидно и вытекает из единственности аффинного преобразования, так как векторы y_i в данном случае образую базис пространства $\mathbb{R}^3$. Рассмотрим второй случай. Пусть векторы y_i лежат в одной плоскости. Тогда существуют такие $\alpha, \beta \in \mathbb{R}$, что

$$y_3 = \alpha y_1 + \beta y_2$$

(заметим, что $\alpha + \beta \neq 1$, т.к. точки не лежат на одной прямой). Тогда, с одной стороны, с учетом этого равенства, из условий леммы для $i = 1, 2$ вытекает

$$Cy_3 = C(\alpha y_1 + \beta y_2) = \alpha C y_1 + \beta C y_2 =$$
$$= \alpha(\tilde{C}y_1 + P) + \beta(\tilde{C}y_2 + P) = \tilde{C}y_3 + (\alpha + \beta)P.$$

С другой стороны, из условий леммы для $i = 3$ следует, что

$$Cy_3 = \tilde{C}y_3 + P.$$

Таким образом выполняется

$$\tilde{C}y_3 + (\alpha + \beta)P = Cy_3 = \tilde{C}y_3 + P,$$

что, с учетом соотношения $\alpha + \beta \neq 1$ возможно лишь в случае $P = 0$. В этом случае соотношения из условий леммы принимают вид

$$Cy_i = \tilde{C}y_i, i = 1..3,$$

откуда с очевидностью следует равенство $C = \tilde{C}$. Лемма 1 доказана.

Лемма 2. В условиях утверждения 1 существует такой линейны оператор $V \in SO(3)$, что

$${}^2a_i = V({}^1a_i), \ \forall i = 1..3.$$

Доказательство: Аналогично лемме 1 рассмотрим 2 случая

1) векторы ${}^1a_i$ не лежат в одной плоскости;

2) векторы ${}^1a_i$ лежат в одной плоскости.

Случай 1: векторы ${}^1a_i$ не лежат в одной плоскости, а значит образуют базис пространства $\mathbb{R}^3$. В этом случае из условия (4) следует, что векторы ${}^2a_i$ также образуют базис пространства $\mathbb{R}^3$. Хорошо известно, что в данном случае существует единственный линейны оператор $V : \mathbb{R}^3 \mapsto \mathbb{R}^3$ такой, что

$$V({}^1x - {}^1y_i) = {}^2x - {}^2y_i, \ i = 1..3.$$

Т.к. множество $\{y_i, i = 1..3\}$ подчинено одному аффинному преобразованию, то очевидно, что

$$\|{}^2a_i - {}^2a_j\| = \|{}^2y_i - {}^2y_j\| = \|{}^1y_i - {}^1y_j\| = \|{}^1a_i - {}^1a_j\| \ 1 \leqslant i, j \leqslant 3,$$

откуда, при учете условия 3, вытекает, что $\|Vz\| = \|z\| \ \forall z \in \mathbb{R}^3$, т.е. K - изометрический оператор. Тогда из условия (4) мгновенно следует, что $detV = 1$, а значит $V \in SO(3)$. Таким образом лемма доказана для 1го случая.

Случай 2: векторы ${}^1x - {}^1y_i$ лежат в одной плоскости. Заметим сначала, что из условия 3 следует, что существуют такие $\alpha, \beta \in \mathbb{R}$, что

$$ {}^ka_3 = \alpha{}^ka_1 + \beta{}^ka_2, \ k = 1, 2. \tag{5}$$

Положим ${}^kz = {}^ka_1 \otimes {}^ka_2, \ k = 1, 2$. Тогда для векторов ${}^kz, {}^ka_1, {}^ka_2$ выполняются условия случая 1, а значит существует такой линейный оператор $V \in SO(3)$, что

$${}^2x - {}^2y_i = V({}^1x - {}^1y_i), \ \forall i = 1..2.$$

Тогда из линейности V и (5) следует, что

$${}^2x - {}^2y_i = V({}^1x - {}^1y_i), \ \forall i = 1..3.$$

Лемма 2 полностью доказана.

Доказательство УТВЕРЖДЕНИЯ 1: По лемме 2 существует такой линейны оператор $V \in SO(3)$, что

$${}^2a_i = V({}^1a_i), \ \forall i = 1..3.$$

С другой стороны по условиям утверждения $\exists C \in SO(3), T \in \mathbb{R}^3 :: {}^2y_i = C{}^1y_i + T \ \forall i = 1..3)$. Тогда предыдущее равенство принимает вид

$${}^2x - C{}^1y_i - T = V({}^1x - {}^1y_i) \Rightarrow$$

$$V({}^1y_i) = C{}^1y_i + T - {}^2x + V{}^1x = C{}^1y_i + d,$$

где $d = T - {}^2x + V\,{}^1x$. Тогда по лемме 1 $V = C$ и $d = 0$, т.е.

$${}^2x = C\,{}^1x + T$$

Утверждение 1 доказано.

Замечание. *Вполне очевидно, что условия, указанные в утверждении 1 являются также необходимыми, т. е. утверждение 1 в полной мере является критерием. В дальнейшем, в обозначениях утверждения 1, множество точек, подчиненных тому же аффинному преобразованию, что и точки y_1, y_2, y_3 для моментов времени t_1, t_2 , будем обозначать как $AffCl_{t_1}^{t_2}(y_1, y_2, y_3)$.*

С помощью данного критерия мы можем модифицировать классическое применение схемы RANSAC следующим образом:

1)на k-ой итерации случайным образом выбираются 3 точки $y_{1,k}, y_{2,k}$, $y_{3,k}$, не лежащие на одной прямой, удовлетворяющие условию $\|{}^1y_{i,k} - {}^1y_{j,k}\| = \|{}^2y_{i,k} - {}^2y_{j,k}\|,\ i, j = 1..3$;

2)для всех остальных точек в соответствии с утверждением 1 проверяется подчиненность их тому же аффинному преобразованию;

3)если количество подходящих точек $\left|AffCl_{t_1}^{t_2}(y_{1,k}, y_{2,k}, y_{3,k})\right|$ больше, чем на предыдущей итерации, то текущий набор подходящих точек полагается "инлаерами";

4)для следующей итерации выбираются точки $y_{1,k+1} y_{2,k+1}, y_{3,k+1}$, удовлетворяющие условию из шага 1, так, чтобы $\{y_{1,k+1}, y_{2,k+1}, y_{3,k+1}\} \not\subseteq AffCl_{t_1}^{t_2}(y_{1,m}, y_{2,m}, y_{3,m})\ \forall 1 \leqslant m \leqslant k$;

5)Количество итераций может определяться тем же правилом, что и в классической версии RANSAC схемы.

Условие в шаге 4 обусловлено тем, что $AffCl_{t_1}^{t_2}(y_{1,k}, y_{2,k}, y_{3,k}) =$
$= AffCl_{t_1}^{t_2}(z_1, z_2, z_3)$ для любых $z_1, z_2, z_3 \in AffCl_{t_1}^{t_2}(y_{1,k}, y_{2,k}, y_{3,k})$, удовлетворяющих условию, указанному в шаге 1.

5 Заключение

Ключевым пунктом данной статьи является п.4, в котором был сформулирован критерий, который можно использовать вместо шагов 2 и 3 в схеме RANSAC в п.3. Однако не были рассмотрены возможные модификации для условий остановки итерационного процесса RANSAC, что может оказаться перспективным в ключе данной статьи. Также не был проведен сравнительный анализ данной модификации с классической схемой, что также представляется перспективным направлением для выявления достоинств и недостатков предложенного алгоритма по сравнению с алгоритмом, описанным в п.3.

ЛИТЕРАТУРА

[1] Е.А. Девятериков, Б.Б.Михайлов ВИЗУАЛЬНЫЙ ОДОМЕТР // Вестник МГТУ им. Н.Э. Баумана. Сер. “Приборостроение”. 2012 с. 68-82

[2] Hatem ALISMAIL, Brett BROWNING and M. Bernardine DIAS Evaluating Pose Estimation Methods for Stereo Visual Odometry on Robots

[3] David Nister, Oleg Naroditsky, James Bergen Visual Odometry for Ground Vehical Application //Journal of Field Robotics 23(1), 3–20 (2006)

[4] Xiao-Shan Gao, Xiao-Rong Hou, Jianliang Tang, Hang-Fei Cheng Complete Solution Classification for the Perspective-Three-Point Problem // IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 25, NO. 8, AUGUST 2003 p. 930-943

[5] Tong Ke, Stergios Roumeliotis An Efficient Algebraic Solution to the Perspective-Three-Point Problem

[6] Joel A. Hesch and Stergios I. Roumeliotis A DirectLeast-Squares(DLS)MethodforPnP // 2011 IEEE International Conference on Computer Vision p. 383-390

[7] Martin A. Fischler and Robert C. Bolles Random Sample Consensus: A Paradigm for Model Fitting with Apphcatlons to Image Analysis and Automated Cartography // Communications of the ACM June 1981, Volume 24, Number 6 p. 381-395