УДК 519.853.4

ИДЕНТИФИКАЦИЯ СЛУЧАЙНЫХ ОТОБРАЖЕНИЙ ТОЧЕЧНЫХ МНОЖЕСТВ В КОНЕЧНОМЕРНЫХ ПРОСТРАНСТВАХ

О. А. Лавягина[1, a]

[1] *Национальный технический университет Украины «Киевский политехнический институт имени Игоря Сикорского», Физико-технический институт*

Аннотация

Одной из актуальных проблем компьютерного зрения является оценка положения камеры по облакам точек (или точечным множествам). В данной работе проводится обзор постановки задачи оценки положения камеры по двум облакам точек, а также изучается алгоритм сопоставления двух точечных множеств как решение поставленной задачи.

Ключевые слова: метод наименьших квадратов, сингулярное разложение матрицы, трёхмерные точечные множества, сканирование трёхмерных объектов

Введение

Оценка положения камеры по облакам точек (или точечным множествам) лежит в основе сканирования объектов с помощью 3D сканера, а также одновременной локализации и картографирования. Для решения этих задач используется итеративный алгоритм ближайших точек и его модификации. В связи с развитием и компактизацией вычислительной техники появилась возможность реализовывать алгоритм на маленьких компьютерах (например, бортовые компьютеры дронов). Поскольку нужно, чтобы такие устройства работали надёжно, было решено тщательно изучить алгоритм, который используется. Таким образом, целью данной работы является исследование новых свойств сходимости итеративного алгоритма ближайших точек как метода оценки положения камеры по облакам точек.

1. Постановка задачи

Есть два множества: исходное $S \subset \mathbb{R}^3$ (source) и целевое $T \subset \mathbb{R}^3$ (target). Точки исходного множества $\boldsymbol{s} \in S$ повернули с помошью матрицы поворота $R \in \mathbb{R}^{3\times 3}$, $R^T = R^{-1}$, $\det R = 1$ и сдвинули с помощью вектора $\boldsymbol{b} \in \mathbb{R}^3$. Также в процессе сканирования появился аддитивный гауссовский шум с неизвестной дисперсией $\boldsymbol{k_s} = R \cdot \boldsymbol{s} + \boldsymbol{b} + \boldsymbol{\xi_s}$, $\boldsymbol{\xi_s} \sim N\left(0, \sigma^2 \cdot I\right)$, где $k \ : \ S \to T$ – разметка, то есть функция, которая сопоставляет каждую точку исходного множества с точкой целевого множества.

Задача состоит в таком выборе матрицы R и вектора $\boldsymbol{b}$, при которых расстояние между $\boldsymbol{k_s}$ и $R \cdot s + \boldsymbol{b}$ для всех $\boldsymbol{s} \in S$ было бы наименьшим

$$\sum_{\boldsymbol{s}\in S} \|\boldsymbol{k_s} - R \cdot \boldsymbol{s} - \boldsymbol{b}\|^2 \to \min_{k,R,\boldsymbol{b}} . \tag{1}$$

[a] olya.lavyagina@gmail.com

2. Применение метода наименьших квадратов

Когда множества S и T конечные, для решения поставленной задачи можно воспользоваться обычным методом наименьших квадратов [1]. Для этого требуем, чтобы сумма квадратов погрешностей достигала минимума. Другими словами,

$$\sum_{\boldsymbol{s}\in S} \|\boldsymbol{\xi_s}\|^2 = \sum_{\boldsymbol{s}\in S} \|\boldsymbol{k_s} - R \cdot \boldsymbol{s} - \boldsymbol{b}\|^2 \to \min_{k,R,\boldsymbol{b}} .$$

Сумма квадратов отклонений между векторами – это то же самое, что и сумма квадратов отклонений между проекциями по каждой координате. Множества параметров, входящих в каждую из трёх сумм, разные, поэтому можем минимизировать суммы квадратов проекций отклонений на каждую координату отдельно.

Запишем вектор параметров для координаты x. Он имеет вид

$$\boldsymbol{\theta}^T = \begin{bmatrix} b_x & r_{xx} & r_{xy} & r_{xz} \end{bmatrix} .$$

Тогда МНК-оценка имеет вид $\hat{\boldsymbol{\theta}} = \left(A^T \cdot A\right)^{-1} \cdot A^T \cdot \boldsymbol{k}_x$, где A – конструкционная матрица, первый столбец которой единичный, а по строкам записаны векторы $\boldsymbol{s}$. Для остальных координат оценка вектора параметров имеет аналогичный вид. При этом оценка $\hat{\boldsymbol{\theta}}$ является несмещённой оценкой параметра $\boldsymbol{\theta}$, обладающая наименьшей дисперсией среди класса несмещённых оценок, которые представляют собой линейные комбинации исходных данных $\boldsymbol{k}_x$.

Не смотря на хорошие свойства полученных оценок матрицы R и вектора b, в решении не учитываются ограничения, а именно: матрица поворота R должна быть оргональной, а её определитель должен равняться единице. Также с помощью метода наименьших квадратов мы не можем найти опимальную разметку k, потому что это дискретная функция.

3. Итеративный алгоритм ближайших точек

Итеративный алгоритм ближайших точек (Iterative Closest Points, ICP) [2] состоит из двух чередующихся операций. Инициализируется алгоритм единичной матрицей поворота $R = I$ и нулевым вектором смещения $\boldsymbol{b} = \mathbf{0}$.

Первая итерация состоит в поиске такой разметки $k : S \to T$, чтобы

$$\sum_{\boldsymbol{s}\in S} \|R \cdot \boldsymbol{s} + \boldsymbol{b} - \boldsymbol{k_s}\|^2 \to \min_k,$$

где R и $\boldsymbol{b}$ фиксированы. Таким образом, для каждой точки $\boldsymbol{s} \in S$ находим точку $\boldsymbol{t} \in T$ такую, чтобы расстояние между парами $R \cdot \boldsymbol{s} + \boldsymbol{b}$ и $\boldsymbol{t}$ для всех $\boldsymbol{s} \in S$ было наименьшим

$$\|R \cdot \boldsymbol{s} + \boldsymbol{s} - \boldsymbol{t}\|^2 \to \min_{t\in T}.$$

На следующей итерации происходит поиск поворота R и смещения $\boldsymbol{b}$ при текущей разметке

$$\sum_{\boldsymbol{s}\in S} \|R \cdot \boldsymbol{s} + \boldsymbol{b} - \boldsymbol{k_s}\|^2 \to \min_{R,\boldsymbol{b}}.$$

При этом матрица $R \in SO(3)$, то есть ортогональная матрица размерности 3×3 с определителем $+1$, которая в качестве линейного преобразования действует как поворот.

Пусть R фиксирована. Минимизируем функцию

$$E(\boldsymbol{b}) = \sum_{\boldsymbol{s}\in S} \|R \cdot \boldsymbol{s} + \boldsymbol{b} - \boldsymbol{k_s}\|^2.$$

Приравняв производную от E по $\boldsymbol{b}$ к нулю, найдём оптимальное смещение

$$\boldsymbol{b} = \overline{\boldsymbol{k}}_{\boldsymbol{s}} - R \cdot \overline{\boldsymbol{s}}, \tag{2}$$

где

$$\overline{\boldsymbol{s}} = \frac{\sum_{\boldsymbol{s}\in S} \boldsymbol{s}}{|S|}, \ \overline{\boldsymbol{k}}_{\boldsymbol{s}} = \frac{\sum_{\boldsymbol{s}\in S} \boldsymbol{k_s}}{|S|}.$$

Переформулируем задачу так, чтобы смещение было равно нулю. Пусть $\tilde{\boldsymbol{s}} = \boldsymbol{s} - \overline{\boldsymbol{s}}$, $\tilde{\boldsymbol{k}}_{\boldsymbol{s}} = \boldsymbol{k_s} - \overline{\boldsymbol{k}}_{\boldsymbol{s}}$, тогда

$$R = \operatorname*{argmin}_{R\in SO(3)} \sum_{\boldsymbol{s}\in S} \left\|R \cdot \tilde{\boldsymbol{s}} - \tilde{\boldsymbol{k}}_{\boldsymbol{s}}\right\|^2.$$

После упрощения получим

$$R = \operatorname*{argmax}_{R\in SO(3)} tr\left(\tilde{K}^T \cdot R \cdot \tilde{S}\right),$$

где $\tilde{S}$ и $\tilde{K}$ – матрицы размерности $3 \times |S|$ со столбцами $\tilde{\boldsymbol{s}}$ и $\tilde{\boldsymbol{k}}_{\boldsymbol{s}}$ соответственно.

Используя свойство следа квадратной матрицы [3], получим задачу

$$R = \operatorname*{argmax}_{R\in SO(3)} tr\left(R \cdot \tilde{S} \cdot \tilde{K}^T\right).$$

Обозначим $X = \tilde{S} \cdot \tilde{K}^T$ и возьмём сингулярное разложение [4] этой матрицы, то есть $X = U \cdot \Sigma \cdot V^T$, где $U \in \mathbb{R}^{3\times3}$ и $V \in \mathbb{R}^{3\times3}$ – ортогональные матрицы, а $\Sigma \in \mathbb{R}^{3\times3}$ – диагональная матрица с неотрицательными элементами, причём $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq 0$.

Убедимся, что сингулярное разложение в данном случае единственное. Пусть $|S| = n < \infty$. Определитель матрицы X – это многочлен от величин $\tilde{k}^T_{s_i j}$, где $i = \overline{1,n}$, $j \in \{x, y, z\}$. Найдём вероятность того, что этот определитель равен нулю

$$P(\det X = 0) =$$
$$= M\mathbb{1}\left\{F_{\tilde{k}^T_{s_1 y}, \tilde{k}^T_{s_1 z}, \ldots, \tilde{k}^T_{s_n x}, \tilde{k}^T_{s_n y}, \tilde{k}^T_{s_n z}}\left(\tilde{k}_{s_1 x}\right) = 0\right\},$$

где $F_{a,b,c,\ldots,d}(\cdot)$ – многочлен. Фиксируем элементы, которые стоят в его индексе. Так как $\tilde{k}_{s_1 x}$ – случайная величина с непрерывным распределением, то она принимает фиксированное значение с нулевой вероятностью. Таким образом, $P\left\{F_{a,b,c,\ldots,d}\left(\tilde{k}_{s_1 x}\right) = 0\right\} = 0$, а значит и $P(\det X = 0) = 0$, то есть имеет место неравенство $\sigma_1 \geq \sigma_2 \geq \sigma_3 > 0$. Это означает, что сингулярное разложение матрицы X будет единственным с точностью до перестановки знаков элементов в матрицах U и V.

Подставим сингулярное разложение в след

$$tr\left(R \cdot \tilde{S} \cdot \tilde{K}^T\right) = tr(R \cdot X) = tr\left(R \cdot U \cdot \Sigma \cdot V^T\right) =$$
$$= tr\left(\Sigma \cdot V^T \cdot R \cdot U\right).$$

Заметим, что V, R и U – ортогональные матрицы, поэтому матрица $M = V^T \cdot R \cdot U$ также ортогональная. Это означает, что $\boldsymbol{m}_i \cdot \boldsymbol{m}_i^T = 1$ для каждой строки $\boldsymbol{m}_i$ матрицы M. Следовательно, модули всех элементов m_{ij} матрицы M не превосходят единицы. Вспомним, что Σ – диагональная матрица с положительными элементами. Поэтому

$$tr(\Sigma \cdot M) = \sum_{i=1}^{3} \sigma_i \cdot m_{ii} \leq \sum_{i=1}^{3} \sigma_i.$$

Таким образом, след максимизируется при

$$m_{ii} = 1 \, \forall i = \overline{1,3}.$$

Так как M – ортогональная матрица, то она должна быть единичной $I = M = V^T \cdot R \cdot U$, откуда $R = V \cdot U^T$.

Заметим, что сейчас R – это ортогональная матрица, но при этом возможны две ситуации: когда $\det R = 1$, то есть матрица R действует как поворот, и $\det R = -1$, то есть матрица R действует как поворот и отражение. $\det V \cdot U^T = 1$ эквивалентно тому, что $\det M = 1$. По теореме, доказанной Альфредом Хорном [5], M – матрица поворота тогда и только тогда, когда её диагональ (m_{11}, m_{22}, m_{33}) лежит в выпуклой оболочке точек $(\pm1, \pm1, \pm1)$, где чётное число координат (в том числе и ни одной) равно -1. Чтобы противоречить этому условию, оптимальная диагональ матрицы M должна иметь вид $(1, 1, -1)$, так как m_{33} – это коэффициент при наименьшем сингулярном числе σ_3.

Таким образом, искомая матрица поворота R имеет вид [6]

$$R = V \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \det\left(V \cdot U^T\right) \end{bmatrix} \cdot U^T,$$

а оптимальный вектор сдвига вычисляется по формуле (2).

3.1. Сходимость итеративного алгоритма ближайших точек

Покажем, что итеративный алгоритм ближайших точек всегда сходится. Предположим, что алгоритм никуда не сходится. Тогда существует такое конечное множество S, матрица поворота R, вектор смещения $\boldsymbol{b}$ и шум $\boldsymbol{\xi_s}$, что алгоритм будет выполнять бесконечно много итераций и никогда не остановится.

Пусть алгоритм перебрал все возможные разметки, то есть выполнил $|S|^{|S|}$ итераций, и не остановился. В этом случае на следующем шаге будет выбрана разметка, которая уже выбиралась на одном из предыдущих шагов. Тогда выражение (1) либо увеличится, либо не изменится. Первый случай невозможен. Во втором случае алгоритм завершит свою работу. Это означает, что алгоритм всегда сходится.

При этом алгоритм имеет ограниченное применение и хорошо работает только в том случае, если шум, а также начальный поворот и сдвиг множества S относительно множества T малы. На рис. 1 изображён пример двух множеств, для которых алгоритм не даёт оптимального результата.

Множества представляют собой два идентичных прямоугольних треугольника с катетами, равными единице, которые отличаются углом поворота π и сдвигом

$$\boldsymbol{b} = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}^T.$$

Множество T изображено пунктиром. Это треугольник с вершинами $(0,0), (0,1)$ и $(1,0)$.

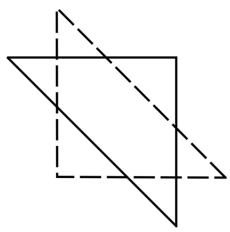

Рис. 1. Множества, для которых ICP не даёт оптимального результата

Полученные с помощью алгоритма взаимные размещения множеств изображены на рис. 2.

Правильным результатом были бы такие поворот и смещение, когда все соответствующие точки двух треугольников имели бы одинаковое положение. Из-за того, что точки не являются соответсвующими, глобальный минимум не достигается.

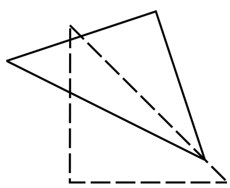

Рис. 2. Результат ICP для треугольников

4. Полный перебор вершин

Оптимальное решение поставленной задачи можно найти полным перебором: рассматриваем все возможные разметки. Для каждой из них применяем второй шаг итеративного алгоритма ближайших точек. Находим поворот и смещение, а затем выбираем ту разметку, которая минимизирует расстояние между множествами. В данном случае будет найден глобальный минимум, однако метод вычислительно неэффективен, так как его сложность – n^n, где n – количество вершин в исходном множестве.

Выводы

Задача оценки положения камеры на сегодняшний день не решена, а потому её существующие решения не для всех входных данных дают оптимальный результат. Вычислительная сложность некоторых процедур, а также наложенные в задаче ограничения, не позволяют построить корректный алгоритм решения, что приводит к неправильным результатам.

Данная работа содержит постановку этой задачи, а также её решение итеративным алгоритмом ближайших точек. Исследование определителя матрицы, для которой применяется сингулярное разложение, показало, что это разложение единственно. Это значит, что алгоритм описан однозначно. Не смотря на то, что алгоритм не всегда даёт оптимальный результат, он сходится, в то время как решение полным перебором может занять много времени.

Перечень использованных источников

1. Luenberger D.G. Optimization by Vector Space Methods. Professional Series. — Wiley, 1997. — ISBN: 9780471181170.
2. Zhang Zhengyou. Iterative point matching for registration of free-form curves and surfaces // International journal of computer vision. — 1994. — October. — Vol. 13. — P. 119–152.
3. Fukunaga K. Introduction to Statistical Pattern Recognition. Computer science and scientific computing. — Elsevier Science, 2013. — ISBN: 9780080478654.
4. Golub Gene H., Loan Van, F. Charles. Matrix Computations. — Third edition. — The Johns Hopkins University Press, 1996.
5. Horn Alfred. Doubly Stochastic Matrices and the Diagonal of a Rotation Matrix // American Journal of Mathematics. — 1954. — July. — Vol. 76. — P. 620–630.
6. Sorkine-Hornung Olga, Rabinovich Michael. Least-Squares Rigid Motion Using SVD. — 2017. — January. — Technical note.