

# ИССЛЕДОВАНИЕ АЛГОРИТМА ИЗВЛЕЧЕНИЯ ТРЕХМЕРНОЙ СТРУКТУРЫ ОБЪЕКТОВ ИЗ ИХ ФОТОИЗОБРАЖЕНИЙ ДЛЯ РЕКОНСТРУКЦИИ ГЕОМЕТРИЧЕСКИХ МОДЕЛЕЙ СЦЕН ГОРОДСКИХ ЛАНДШАФТОВ

Ковальский С.В., Зори С.А.

Кафедра ПМИ ДонНТУ

skovalsky@yandex.ru, zori@pmi.dgtu.donetsk.ua

## **Abstract**

*Kovalsky S.V., Zori S.A. Research of algorithm of 3-dimensional object structure extraction from their photos for geometric models reconstruction of city landscape scenes. Algorithm of 3-dimensional structure extraction from set of straight lines given on the set of images is described. Formulation of local and global minimization procedures is given.*

## **1. Обзор состояния решения задачи СИД**

В настоящее время широкое применение находят системы компьютерной графики, формирующие изображения на основе некоторой трехмерной модели – тренажеры транспортных средств, авиасимуляторы, навигационные системы и т.п. Для эксплуатации таких систем необходимо представление виртуальной или реальной сцены в виде ее трехмерной модели, описывающей ее геометрию, визуальные свойства и т.п.. При симуляции (построении высокореалистичных динамических наборов изображений) реальных сцен и объектов, отличающихся высокой степенью сложности, неоднородности и геометрической нетривиальностью, таких как природные или городские ландшафты, информация о трехмерной геометрической модели чаще всего недоступна, а визуальные свойства сцен характеризуют их фотографические изображения. Данная задача впервые была поставлена в [1] и получила название “Структура из движения” – СИД (“Structure-from-Motion” – SFM). Существующие алгоритмы СИД строят трехмерную структуру на основании двухмерных проекций опорных элементов, размеченных на фотографических изображениях.

Классификация алгоритмов СИД была подробно рассмотрена авторами в [2].

Для решения этой задачи в некоторых подходах сцена, структуру которой требуется определить, представляется в виде набора опорных точек, данных на множестве плоских изображений. В [3, 4] показано, что структура (взаимное расположение) множества точек может быть

извлечена из двух изображений даже в случае, если относительное расположение двух камер неизвестно.

Некоторые исследователи [5, 6] ставят задачу извлечения структуры сцены, состоящей из отрезков прямых. При решении задачи СИД для реконструкции городских ландшафтов предпочтительно использование именно этого типа опорных элементов, как основного в структуре реконструируемых сцен.

При построении алгоритмов СИД используется ограниченный ряд математических объектов, таких как матрицы поворота, единичные векторы, ортогональные векторы, для которых нет однозначного параметрического представления. Поскольку в настоящее время не определены лучшие способы представления перечисленных математических объектов, то задачу выбора лучшего способа их параметризации следует считать нерешенной.

В данной работе выполняется анализ глобальной и локальной процедур минимизации алгоритма извлечения трехмерной структуры из набора отрезков прямых линий, предлагаемого в [5], как одного из самых трудоемких этапов решения задачи СИД, в целях определения лучшего способа его реализации и возможных модификаций, повышающих временные характеристики или качество работы алгоритма. Проводится также исследование возможных способов представления используемых алгоритмом математических объектов.

Алгоритм [5] выполняет реконструкцию трехмерной геометрической модели сцены, описанной исходными данными. Результатом работы алгоритма является набор прямых в пространстве, соответствующих опорным элементам, а также описание положений камеры при получении набора исходных кадров. Это дает возможность в дальнейшем получить традиционное описание сцены в виде набора вершин, ребер и граней.

Исходными данными для алгоритма являются координаты в экранной плоскости конечных точек отрезков прямых линий, которые являются проекциями выбранных опорных элементов – ребер трехмерных объектов. Алгоритм был построен на основе перспективной модели проецирования и при условии известности внутренних параметров камеры, что определило использование фокусного расстояния камеры в качестве единицы измерения расстояния в экранной плоскости.

## **2. Анализ алгоритма реконструкции**

Подробный анализ алгоритма показал, что он основан на процессе минимизации функции  $O$ , который дает оценку отклонения реконструируемой модели, позволяющей построить расчетные проекции опорных элементов, от имеющихся исходных данных, представленных фактическими проекциями опорных элементов. Для выполнения

минимизации функции  $O$  важно выбрать наиболее подходящий метод, т.к. методы минимизации характеризуются различной степенью чувствительности к форме поверхности минимизируемой функции и, соответственно, различной скоростью сходимости и временем работы. Эти характеристики являются чрезвычайно важными в свете обработки в режиме реального времени, мощности применяемых средств вычислительной техники и качества реконструирования реальных сцен.

Функция  $O$  является нелинейной и может иметь несколько локальных минимумов. Большинство методов минимизации нелинейных функций сходится к тому или иному локальному минимуму в зависимости от выбора начальных значений параметров минимизации. Поэтому авторы [5] предлагают генерировать наборы начальных значений параметров минимизации случайным образом и искать соответствующие локальные минимумы функции  $O$  до тех пор, пока не выполнится некоторое условие, позволяющее считать очередной локальный минимум глобальным.

Функция  $O$  зависит от параметров  $R_j$ ,  $t_j$ ,  $\hat{v}_i$ ,  $d_i$ , однако в [5] предлагается механизм предварительной оценки параметров минимизации  $t_j$ ,  $\hat{v}_i$ ,  $d_i$  на основе значений параметров  $R_j$ , что позволяет в процессе глобальной минимизации генерировать случайным образом только параметры  $R_j$ .

После генерирования значений параметров  $R_j$  выполняются 4 этапа локальной минимизации, первые 3 из которых предназначены для оценки начальных значений параметров  $t_j$ ,  $\hat{v}_i$  и  $d_i$ , а 4-й представляет собой минимизацию функции  $O$  (табл.1).

Таблица 1 – Характеристики этапов локальной минимизации

№ этап а	Назначение и способ минимизации	Параметры минимизации	Количество степеней свободы
1	Получение начальной оценки параметров $\hat{v}_i$ путем минимизации функции $C_1$	$\hat{v}_i$	$2n$
2	Улучшение оценок параметров $R_j$ и $\hat{v}_i$ путем минимизации функции $C_1$	$R_j, \hat{v}_i$	$2n + 3(m - 1)$
3	Получение начальной оценки параметров $t_j$ и $d_i$ путем минимизации функции $C_2$	$t_j, d_i$	$2n + 3(m - 1) - 1$
4	Минимизация функции $O$ с использованием оценок всех параметров, полученных предыдущими этапами	$R_j, t_j, \hat{v}_i, d_i$	$4n + 6(m - 1) - 1$

Из таблицы 1 видно, что этапы 1-3 локальной минимизации имеют значительно меньшее по сравнению с 4-м этапом количество степеней

свободы, что дает основание считать их относительно менее сложными с вычислительной точки зрения.

На 1-м и 2-м этапах минимизируется функция  $C_1$ :

$$C_1 = \sum_{j=1}^m \sum_{i=1}^n (\hat{m}_{ij}^T R_j \hat{v}_i)^2 \quad (2)$$

где  ${}^c \hat{m}'$  – нормаль к плоскости, проходящей через точку фокуса камеры и проекцию опорной линии на данный кадр, которую можно рассчитать, выполнив векторное умножение векторов, соединяющих точку фокуса с концами проекции опорной линии (см. рис. 1):

$$\begin{aligned} {}^c m' &= (x_1 \ y_1 \ 1) \times (x_2 \ y_2 \ 1) \\ {}^c \hat{m}' &= {}^c m' / \| {}^c m' \| \end{aligned} \quad (3)$$

Следует отметить, что в [5] для расчета  ${}^c \hat{m}'$  авторы приводили формулу:

$${}^c m' = (x_1 \ y_1 - 1) \times (x_2 \ y_2 - 1), \quad (4)$$

которая не соответствует схеме проецирования, приводимой в [5], т.к. предполагает направленность оси Z от экранной плоскости в сторону точки фокуса. Данное несоответствие позволяет считать формулу (4) ошибочной.

На 3-м этапе минимизируется функция  $C_2$ :

$$C_2 = \sum_{j=1}^m \sum_{i=1}^n (\hat{m}_{ij}^T R_j (d_i - t_j))^2 \quad (5)$$

Для решения задачи минимизации на каждом этапе необходимо построить вектор переменных, на основании которого будет вычисляться значение целевой функции минимизации, и, возможно, вектор ограничений, накладываемых на вектор переменных.

Способ построения вектора переменных, а также наличие (и способ построения при наличии) вектора ограничений определяются способами представления параметров минимизации  $R_j$ ,  $t_j$ ,  $|\hat{v}_i|$ ,  $d_i$ . Отсутствие или наличие ограничений определяет возможность использования различных методов минимизации.

Методы минимизации при отсутствии ограничений делятся на классы 0-го, 1-го или 2-го порядков [7]. Методы 0-го порядка характеризуются относительной простотой реализации, т.к. требуют вычисления значений только собственно целевой функции [7]. Методы 1-го и 2-го порядков более сложны для реализации и требуют большего количества вычислений на каждой итерации, т.к. требуют вычисления значений первой (методы 1-го и 2-го порядков) и второй (2-го порядка) частных производных целевой функции. Следует отметить, что, как правило, методы 1-го и 2-го порядков находят минимум с заданной точностью за значительно меньшее количество итераций, чем методы 0-го

порядка, а при первом взгляде на процессы расчета градиента и матрицы Гессе (матрицы вторых частных производных), выполняемые на каждой итерации, определяется возможность распараллеливания. Это, однако, требует дополнительных исследований и будет выполнено в дальнейшем.

Решение задачи минимизации с ограничениями сводится, как правило, к решению последовательности задач минимизации без ограничений [8]. Т.о., при реализации методов минимизации с ограничениями (например, методов штрафных функций) используются методы минимизации без ограничений. Постановка отдельной задачи в последовательности задач минимизации без ограничений сводится к построению целевой функции, не имеющей ограничений на значения параметров, на основании целевой функции полной задачи и ограничений на значения ее параметров. В зависимости от класса используемого метода минимизации без ограничений, способы реализации методов минимизации с ограничениями можно классифицировать на способы 0-го, 1-го и 2-го порядков. При использовании способа 0-го порядка объем вычислений на каждой итерации решения каждой задачи минимизации без ограничений соответствует общему объему вычислений значения целевой функции полной задачи и ограничений на значения ее параметров.

Одним из факторов, определяющих время нахождения минимума с заданной точностью некоторым методом минимизации, является объем вычислений, выполняемых на каждой итерации, который в свою очередь определяется способом представления параметров минимизации.

Авторами данной работы предлагается рассматривать различные способы представления параметров минимизации  $R_j$ ,  $t_j$ ,  $\hat{v}_j$ ,  $d_i$  в разбиении на группы способов, включающих ограничения на значения переменных и не включающих таковых. При этом в каждой группе становится возможным определить лучший или лучшие способы по критерию минимизации объема вычислений, выполняемых на каждой итерации методами одного порядка. Сравнение объемов вычислений, необходимых для расчета значений целевой функции, позволит определить лучший способ представления параметров минимизации для методов 0-го порядка в группе способов, не включающих ограничений. При определении лучшего способа в группе способов, включающих ограничения, необходимо также сравнивать объемы вычислений, которые выполняются при определении значений ограничений. При этом общая оценка объема вычислений на каждой итерации может быть получена сложением объемов вычислений, выполняемых при расчете значения целевой функции и при расчете значений ограничений. Предлагается также дать приблизительную сравнительную оценку лучших способов групп, сравнив полученные для них оценки объемов вычислений на каждой итерации.

Параметр  $R_j$  описывает ориентацию камеры в пространстве при получении  $j$ -го исходного кадра и представляет собой матрицу вращения в

трехмерном пространстве. Матрица вращения в трехмерном пространстве является квадратной матрицей размерами  $3 \times 3$  и состоит, соответственно, из 9-ти элементов, однако вращение в трехмерном пространстве имеет только 3 степени свободы, т.е. теоретически возможно выразить матрицу вращения через 3 переменных, без использования ограничений. Это характеризует параметр  $R_j$  как наиболее сложный для представления из всех параметров минимизации. Параметр  $t_j$  является трехмерным вектором, не имеет ограничений на значение и поэтому тривиально представляется своими компонентами. Параметр  $\hat{v}_j$  является трехмерным вектором единичной длины и может быть представлен при помощи матрицы вращения [9]. Значение параметра  $d_i$  связано со значением параметра  $\hat{v}_i$ , описание представления  $d_i$  с использованием  $\hat{v}_i$  приводится в [5],[9]. Таким образом, для определения лучшего способа представления параметров минимизации  $R_j$ ,  $\hat{v}_j$ ,  $d_i$  достаточно рассмотреть способы представления вращения в трехмерном пространстве.

Первые пять приведенных в таблице способов представления вращения в трехмерном пространстве описаны в [9], [10], [11]. Оценки количества операций, необходимых для вычисления матрицы вращения, сделаны на основании описания преобразований, приведенных в тех же источниках. Способ преобразованных углов Эйлера является авторским и заключается в следующем: в способе представления углами Эйлера производится замена выражений  $\cos(\alpha), \sin(\alpha), \cos(\beta), \sin(\beta), \cos(\gamma), \sin(\gamma)$ , (где  $\alpha, \beta, \gamma$  – углы Эйлера) новыми переменными, при этом углы Эйлера перестают быть переменными и вводятся ограничения на новые переменные, обеспечивающие выполнение равенств (6).

$$\cos^2(\alpha) + \sin^2(\alpha) = 1; \cos^2(\beta) + \sin^2(\beta) = 1; \cos^2(\gamma) + \sin^2(\gamma) = 1 \quad (6)$$

Способы представления вращения в пространстве углами Эйлера, вектором Гиббса и не единичным кватернионом составляют группу способов, не включающих ограничений. В этой группе лучшим способом представления  $R_j$  по критерию минимизации объема вычислений, выполняемых на каждой итерации методами минимизации 0-го порядка, является представление не единичным кватернионом. Способ представления вращения в пространстве вектором Гиббса требует выполнения незначительно большего количества операций сложения/вычитания, хотя требует введения меньшего количества переменных, чем лучший способ группы. Способ углов Эйлера невыгодно отличается от остальных описанных способов необходимостью применения тригонометрических функций, что требует значительных вычислительных затрат.

Таблица 2 – Сравнительная характеристика различных способов представления  $R_j$ 

№ п / п	Описание способа	Кол- во пере- мен- ных	Кол- во огран- и- чений	Оценка кол-ва операций, необходимых для вычисления R	Оценка кол-ва операций, необходимых для вычисления значений ограничений
1	Углы Эйлера	3	0	1) с большим применением тригонометрических функций: cos, sin: 6 (*): 12 (+),(-): 6 2) с меньшим применением тригонометрических функций: cos, sin: 2 (*): 16 (+),(-): 13	—
2	Вектор Гиббса	3	0	(*): 16 (+),(-): 18 (/): 1	—
3	Параметры Эйлера	4	1	(*): 16 (+),(-): 15	(*): 4 (+),(-): 4
4	Кватернион – не единичный	4	0	(*): 16 (+),(-): 15 (/): 1	—
5	Кватернион –единичный	4	1	(*): 16 (+),(-): 12	(*): 4 (+),(-): 4
6	Преобразова- нные углы Эйлера	6	3	(*): 12 (+),(-): 6	(*): 6 (+),(-): 6

Способы представления вращения параметрами Эйлера, единичным кватернионом и преобразованными углами Эйлера составляют группу способов, включающих ограничения. В этой группе лучшим способом по

критерию минимизации общего количества операций, необходимых для вычисления целевой функций и ограничений, является метод преобразованных углов Эйлера. Однако следует отметить, что методы минимизации с ограничениями требуют дополнительных вычислительных затрат на каждой итерации используемого метода минимизации без ограничений, пропорциональных количеству ограничений. Поэтому в зависимости от выбранного метода минимизации с ограничениями лучшим может оказаться способ представления вращения единичным кватернионом.

Сравнивая лучшие способы групп – представление не единичным кватернионом с представлением преобразованными углами Эйлера, лучшим следует назвать способ представления не единичным кватернионом, как требующий меньшего числа операций умножения/деления и при этом не требующий дополнительных вычислительных затрат, связанных с ограничениями.

На рисунке 1 представлена схема группировки способов представления  $R_j$ .

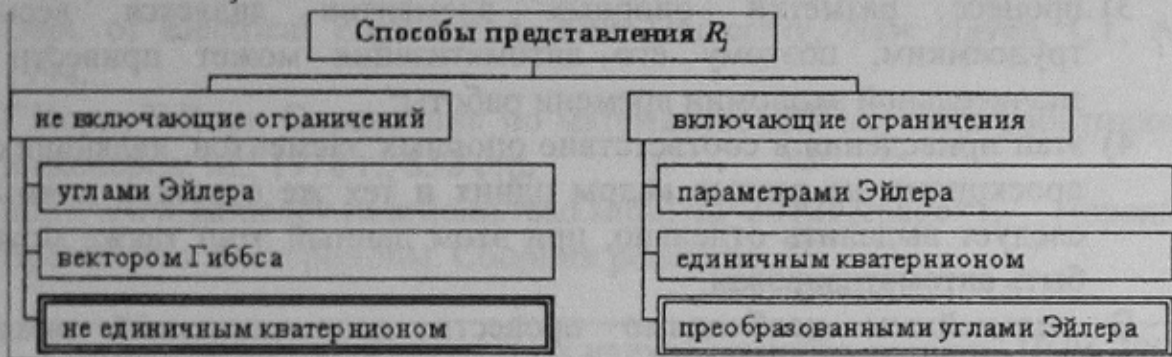


Рисунок 1 – Классификация способов представления  $R_j$

### Заключение

В данной работе проанализирован алгоритм реконструкции трехмерной структуры объектов из их фотоизображений [5] для выполнения реконструкции геометрических моделей сцен городских ландшафтов. В процессе анализа были выделены преимущества и недостатки использования методов минимизации с ограничениями и без ограничений при реализации данного алгоритма, рассмотрены преимущества и недостатки классов методов минимизации без ограничений 0-го, 1-го и 2-го порядков.

Была описана связь между способом представления параметров минимизации и возможностью применения различных методов минимизации. Выполнен выбор лучших способов представления параметров по критерию минимизации объема вычислений, выполняемых на одной итерации методами минимизации без ограничений 0-го порядка и основанными на них методами минимизации с ограничениями.



Были проанализированы целевые функции минимизации, используемые на различных этапах локальной минимизации, и рассмотрены математические объекты, составляющие количество степеней свободы (и соответственно вычислительную сложность) каждого этапа. Сделан вывод о целесообразности выполнения 1-3 этапов локальной минимизации. В процессе верификации математических формул, представленных в [5], была найдена ошибка в (4), приводящая в процессе реализации алгоритма к неадекватным результатам реконструкции.

В процессе построения алгоритма возник ряд задач, которые не были решены в [5] и решение которых предполагается выполнить авторами:

- 1) результаты работы алгоритма включают прямые в пространстве, задача получения отрезков, принадлежащих полученным прямым и соответствующих исходным проекциям, не имеет тривиального решения и должна решаться дополнительно;
- 2) возможно также изменение постановки задачи для получения в качестве конечных результатов не только отрезков, но и более сложных пространственных примитивов;
- 3) процесс разметки опорных элементов является весьма трудоемким, поэтому его автоматизация может привести к значительной экономии времени работы;
- 4) этап приведения в соответствие опорных элементов, являющихся проекциями на разные кадры одних и тех же отрезков прямых, следует выделить отдельно, при этом данный этап также может быть автоматизирован.

В дальнейшем необходимо провести дополнительный анализ возможности применения методов минимизации 1-го и 2-го порядков, в том числе рассмотреть возможность сокращения времени выполнения отдельной итерации метода за счет распараллеливания вычислений. Для работы выбранного метода минимизации могут потребоваться начальные значения переменных минимизации и, следовательно, необходимо решить дополнительную задачу определения этих значений.

### **Литература**

1. S. Ulman, The Interpretation of Visual Motion. The MIT Press, Cambridge, MA, 1979.
2. Ковальский С.В., Зори С.А. Современные алгоритмы извлечения геометрической структуры объектов из фотографических снимков для пространственного моделирования городских ландшафтов. В кн.: Наукові праці Донецького національного технічного університету. Серія "Проблеми моделювання та автоматизації проектування динамічних систем" (МАП-2004). Випуск: 74 - Донецьк: ДонНТУ. - 2004 - 7 с.

3. H.C. Longuer-Higgins, A computer algorithm for reconstructing a scene from two projections. Nature, vol.293, pp.133-135, 1981.
4. J. Weng, T.S. Huang, and N. Ahuja, Motion and structure from two perspective views: Algorithms, error analysis, and error estimation. IEEE Trans. Pattern Anal. Machine Intell., vol.11, no.5, pp.451-476, May 89.
5. C.J. Taylor and D.J. Kriegman. Structure and motion from line segments in multiple images. IEEE Trans. Pattern Anal. Machine Intell., 17(11), November 1995.
6. Olivier D. Faugeras, Francis Lustaman, and Giorgio Toscani, "Motion and structure from point and line matches", in International Conference on Computer Vision, June 1987.
7. [http://matlab.exponenta.ru/optimiz/book\\_2/index.php](http://matlab.exponenta.ru/optimiz/book_2/index.php) А.Г.Трифонов. Постановка задачи оптимизации и численные методы ее решения.
8. [http://matlab.exponenta.ru/optimiz/book\\_1/index.php](http://matlab.exponenta.ru/optimiz/book_1/index.php) А.Г.Трифонов. Optimization Toolbox 2.2 Руководство пользователя.
9. Camillo J. Taylor and David J. Kriegman, "Minimization on the lie group SO(3) and related manifolds", Tech. Rep. 9405, Center for Systems Science, Dept. of Electrical Engineering, Yale University, New Haven, CT, April 1994.
10. Г.Корн, Т.Корн Справочник по математике для научных работников и инженеров. М., 1978 г., 832 стр.
11. <http://www.gamedev.ru/articles/read.shtml?id=30129&page=1>. Норель М. Вращение и кватернионы. Сборник рецептов.

---

Дата надходження до редакції 10.06.2005 р.