

Самокалибрующаяся система трехмерного зрения и анализ ее применимости в робототехнике и спутниковой съемке земного рельефа.

Н.В. Свешникова, А.С. Чернышев, Д.В. Юрин.
ЦОС и ВТ МФТИ, ФГУП НПП ОПТЭКС
Москва, Россия

Аннотация

Разработана расширяемая система восстановления трехмерных сцен, движения и параметров камеры. Путем численного моделирования найден критерий для адаптивного выбора приближения в зависимости от получаемой системой оценки расстояния до объекта. Проведены оценки точности и применимости системы для аэрокосмических задач.

Ключевые слова: Восстановление трехмерных сцен, метод факторизации, триангуляция, VRML модель, комбинированный алгоритм.

1. ВВЕДЕНИЕ

Задача восстановления трехмерных сцен по цифровым изображениям является востребованной и интенсивно развиваемой во всем мире [1-10]. Среди всего комплекса задач выделим три наиболее типичных:

- 1) Восстановление земного рельефа по аэрокосмическим изображениям для нужд картографии и мониторинга промышленных объектов и чрезвычайных ситуаций.
- 2) Восстановление плана помещений по изображениям, полученным с цифровых фотоаппаратов и кинокамер для нужд перепланировки и оборудования промышленных помещений и архитектуры.
- 3) Системы зрения мобильных роботов.

В настоящее время уже есть примеры успешного построения таких систем, однако в целом задача далека от окончательного решения, а часть таких систем требует точных измерений и калибровки, что не всегда удобно. Например, на борту спутника Terra[6] находится два прибора, дающих возможность восстанавливать трехмерные сцены MISR (9 идентичных телескопов с пространственным разрешением 250м) и ASTER (стерео система с пространственным разрешением 15м. Заметим, что при съемке высокого разрешения ($d=1-10м.$) для типичной высоты орбиты $h=1000 км.$ ориентацию камер надо знать с точностью $\sim d/h=10^{-5}-10^{-6}$ рад., а координаты спутника с точностью порядка d . Такие точности в принципе достижимы, но приводит к сложной и дорогостоящей системе со значительными массогабаритными характеристиками. Поэтому представляют интерес подходы, позволяющие восстанавливать эти параметры непосредственно из видео данных.

Большинство описанных в литературе систем, построено на основе какого то одного физического принципа. В то же время, разные подходы [1-4,7-10] взаимно дополняют друг друга и имеют различные «плохие случаи». Таким образом представляется целесообразным построить комбинированную систему, использующую различные алгоритмы восстановле-

ния трехмерных сцен в комбинации, с выбором оптимального для данной сцены и ее различных частей алгоритма.

2. СТРУКТУРА СИСТЕМЫ

В качестве основы системы была выбрана группа алгоритмов, основанных, на факторизации матриц [2-4]. Основанием для такого выбора послужили следующие причины:

- 1) Возможность восстановления не только формы сцены, но и движения, ориентации и фокусных расстояний камер.
- 2) Возможность работы с нестационарными сценами [3].
- 3) Высокая точность метода и из-за больших баз съемки.
- 4) Невысокая вычислительная сложность.

В результате применения подходов [2,4] получаются трехмерные координаты некоторых точек объекта с высокой точностью, по которым может быть построена поверхность в виде сетки или меша [5], координаты и ориентация камер, и при наличии близких объектов фокусные расстояния камер. Структура алгоритмов [2,4] открывает возможности интеграции с другими подходами и сопутствующими данными двумя способами.

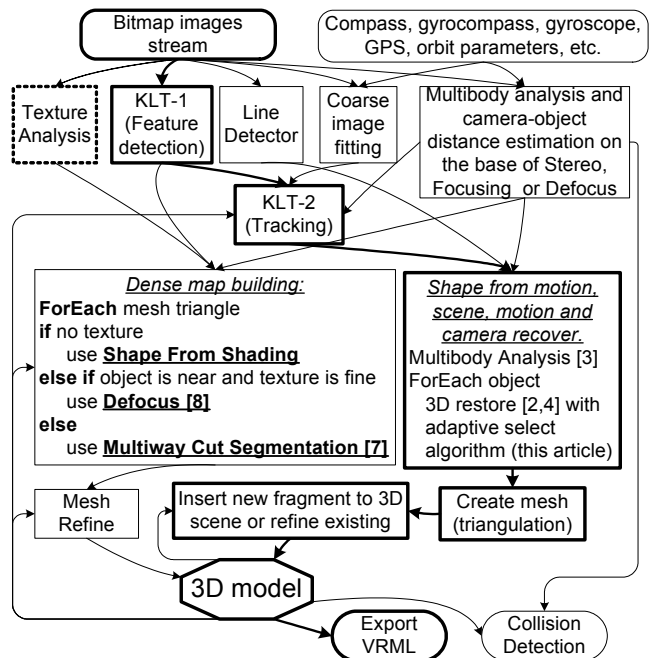


Рис. 1. Общая схема системы.

- 1) Для близких объектов первоначальная оценка расстояний может быть выполнена методами дефокусировки или стерео [8-9], или использованы априорные данные (GPS, гироскопы

и т.п.), что дает информацию для выбора приближения и ускоряет сходимость [4]. Может быть выполнено грубое разделение сцены на области различной удаленности, и для них использованы наилучшие приближения. Использование стерео или дефокусировки с совмещенной камерой [10], позволяет произвести выделение быстро движущихся объектов и облегчить их обработку трэкером характеристических точек.

2) Построенный меш является хорошим начальным приближением с точно определенными узлами. Поверхность может уточняться методами [7-10] до достижения плотной карты глубин сцены. Метод уточнения выбирается на основе анализа текстур изображения и полученного расстояния от камер до уточняемой области объекта, оцениваемой как расстояние от камеры до поверхности, задаваемой мешем.

Общая структура системы изображена на рис.1, реализованные в настоящий момент блоки отмечены жирной рамкой. Система реализована на языке C++ в независимом от операционной системы виде.

3. ОБНАРУЖЕНИЕ И СОПРОВОЖДЕНИЕ ХАРАКТЕРИСТИЧЕСКИХ ТОЧЕК

Идея алгоритма [1] сопровождения состоит в том, что при малых изменениях положения и ориентации камеры, область в более позднем изображении J может быть получена смещением некоторой области предыдущего изображения I . Тогда смещение можно найти из минимизации функционала.

$$\varepsilon = \iint [J(\vec{x} + \frac{\vec{d}}{2}) - I(\vec{x} - \frac{\vec{d}}{2})]^2 w(\vec{x}) d\vec{x} \quad (1)$$

Здесь \vec{d} - вектор смещения, $w(\vec{x})$ - функция окна. Путем линеаризации функций яркости, решение этой задачи сводится к итерационному решению линейной системы [1]:

$$Z\vec{d} = \vec{e} \quad (2)$$

Алгоритм [1] работает с серыми изображениями. Предлагается обобщить подход [1] на случай многоканальных изображений, и, в частности цветных. Это легко достигается заменой в функционале (1) яркостей J, I векторами, компоненты которых соответствуют яркостям в соответствующих каналах. Тогда окончательное уравнение имеет вид такой же, как и уравнение (2), но $Z = \sum_{i=1}^N Z_i$ и $\vec{e} = \sum_{i=1}^N \vec{e}_i$, где Z_i и \vec{e}_i - вычисляются независимо для каждого канала по формулам из [1].

Работа с изображениями ведется по пирамиде разрешений.

4. ФОРМИРОВАНИЕ ПЛОТНОЙ МАТРИЦЫ ИЗМЕРЕНИЙ.

При сопровождении точек по серии изображений, возможна как потеря, так добавление новых отметок. Однако, для алгоритма восстановления трехмерной сцены требуется наличие всех отслеживаемых точек на всех кадрах. Таким образом, появляется задача выделения наибольшей полностью заполненной матрицы, в которой строки - это кадры, а столбцы - отслеживаемые точки.

Предлагается решать эту задачу через поиск клик графа. Граф $\Gamma = \{X, E\}$ строится следующим образом. Множеством

его вершин $X = \{F \cup P\}$ являются все кадры F и все отслеживаемые точки P . Множество ребер E задается двумя условиями. Подмножества F и P являются полностью зависимыми множествами. Ребро между вершинами $f \in F$ и $p \in P$ есть только в том случае, если отметка p присутствует в кадре f , (см. рис. 2).

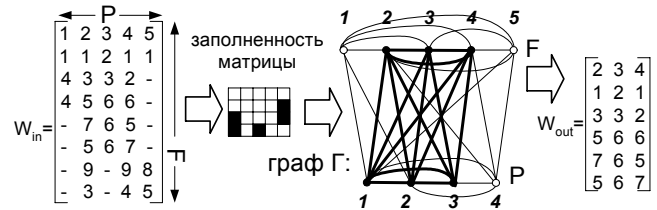


Рис. 2. Поиск плотной подматрицы через клику графа Γ , F, P - число кадров и точек (отметок).

При выделении из графа перекрывающихся клик появляется возможность восстановления отдельных частей сцены с последующей их сшивкой по общим точкам. Это позволяет более полно восстановить сцену, так как алгоритмы [2-4] требуют видимости всех отслеживаемых точек на всех кадрах, что далеко не всегда выполнимо. Для нахождения клики, используется алгоритм [11].

5. АЛГОРИТМ ВОССТАНОВЛЕНИЯ ТРЕХМЕРНОЙ СЦЕНЫ.

В системе реализованы линейные алгоритмы [2] и итерационный алгоритм в перспективной проекции [4]. Заметим, что алгоритм [4] на каждом шагу итерации использует приближение масштабированной ортографии [2], что позволяет в принципе прервать цикл итераций и ограничиться приближенным решением. Так как итерационный алгоритм извлекает дополнительную информацию из перспективных искажений, то при большой дальности до объекта, итерации не приводят к повышению точности. Вопрос о критерии отключения итераций рассматривается в разделе, посвященном результатам.

6. ПОСТРОЕНИЕ ТРЕХМЕРНОЙ СЦЕНЫ В ВИДЕ СЕТКИ, ЭКСПОРТ В VRML.

В результате решения задачи восстановления получаются координаты и ориентация камер, фокусные расстояния (если использовался итерационный алгоритм) и трехмерные координаты тех точек объекта, которые были обнаружены и успешно сопровождалась трэкером (разделы 3-4). Теперь по этим точкам надо построить поверхность объекта в виде меша. В настоящей работе используется следующий подход. Так как в проекции на каждый кадр взаимное расположение отметок всегда правильно, и для кадра с номером f двумерные координаты отметок на этом этапе доступны и содержатся в строках $2f+1$ и $2f+2$ матрицы (3), то триангуляция производится в плоскости кадра в двумерном пространстве по этим точкам. Затем полученные соединения формально применяются к восстановленным трехмерным координатам соответствующих точек. Такой подход позволяет применить быстрый алгоритм [5] и исключает возможность соединения ребром точек с различных поверхностей (сторон) объекта, например при круговой съемке объекта, так как такие точки

пример при круговой съемке объекта, так как такие точки не могут быть видны на изображении одновременно. Восстановленная поверхность, координаты и направления камер могут экспортироваться в формат VRML с натянутыми текстурами.

7. РЕЗУЛЬТАТЫ: АНАЛИЗ ПРИМЕНИМОСТИ И ОЦЕНКА ПОГРЕШНОСТЕЙ.

Предложенный алгоритм тестировался на реальных и синтетических данных. Для синтетических данных в виде трехгранного угла приведены графики сравнительной точности восстановления в различных приближениях в зависимости от отношения расстояния до объекта L к размеру объекта a .

На рис.3-5 видно, что на малых расстояниях точность метода [4] существенно превосходит остальные. Однако по мере увеличения расстояния точность различных приближений становится примерно одинаковой, и для метода [4] начиная с некоторого расстояния становится сопоставима или хуже (рис. 4.5), чем для линейных методов [3]. Это связано с исчезновением перспективных искажений в цифровых изображениях. Поэтому система должна автоматически переключаться на приближение масштабированной ортографии или парapersпективный метод.

Критерий переключения в системе выбирался по результатам численного моделирования на описанных синтетических данных. За основной параметр анализа был принят безразмерный параметр L/a . Очевидно, что точность восстановления зависит и от других параметров:

- углового разрешения камеры;
- углового размера восстанавливаемого объекта;
- отношения максимального перемещения камеры к среднему расстоянию до объекта;
- угла обхода камеры вокруг объекта.

Результаты численного моделирования показывают, что зависимость пороговой величины от этих параметров слабая.

На рис. 3-5 видно, что переключение должно происходить в различных диапазонах для восстанавливаемых величин – формы, ориентации матриц и положения камер: перспективный метод более точно восстанавливает форму в диапазоне расстояний равных $1 - 10^2$, ориентации камер – $1 - 20$, положение камер – $1 - 5$ размеров восстанавливаемого объекта.

На рис. 4. также приведен участок графика, соответствующий съемке объекта в условиях, сопоставимых с условиями спутниковой съемки. Из этого участка графика видно, что приближение масштабированной ортографии и парapersпективное приближение способны давать результат с приемлемой для спутниковой съемки точностью.

На рис. 7-10 представлено восстановление реальной сцены. На рис. 7 и 8 показаны начальный и конечный кадры серии. На рис. 9 отмечены точки, сопровождаемые трекером. Изображение восстановленной VRML-модели представлено на рис. 10. Искажения формы вызваны тем, что на ребрах объекта не были найдены подходящие для сопровождения точки, и алгоритм триангуляции не смог построить реальные границы коробки без них. Однако, координаты узлов меша восстановлены точно. Построенная модель может использоваться как хорошее начальное приближение для уточнения формы объекта.

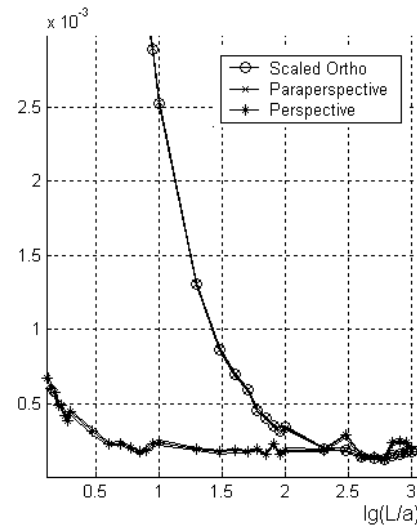


Рис. 3. Погрешность восстановления формы сцены. Угол обхода камеры 60° , разрешение изображения 1024 пикселя.

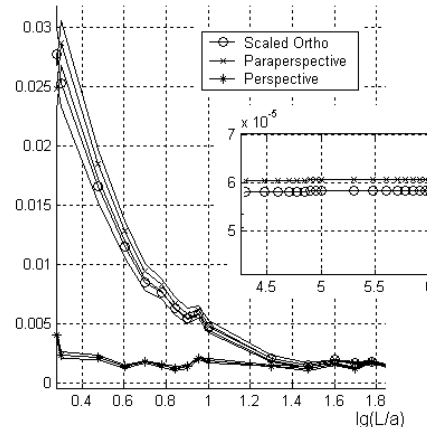


Рис. 4. Погрешность восстановления ориентации камер в радианах. Угол обхода камеры 60° , разрешение изображения 1024 пикселя.

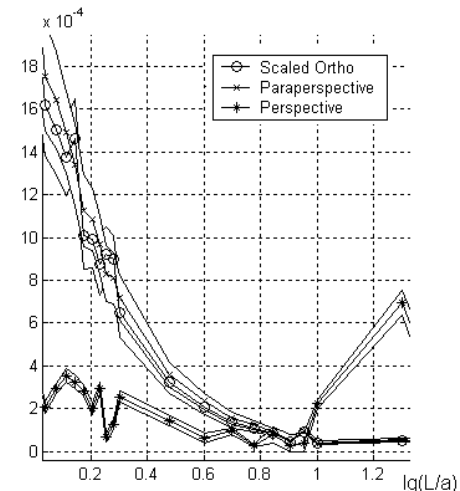


Рис. 5. Погрешность восстановления XY компонент положения камер. Угол обхода камеры 60° , разрешение изображения 1024 пикселя.



Рис. 7.8. Начальный и конечный кадры серии.

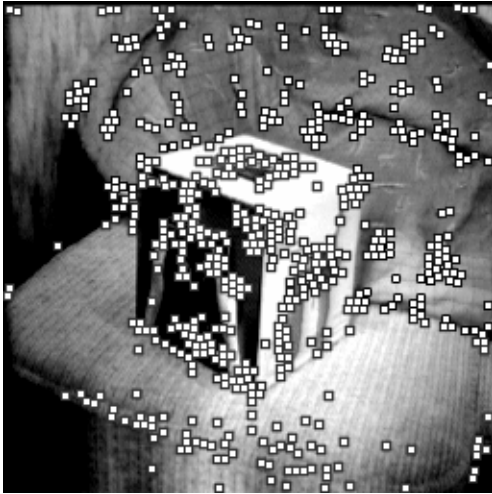


Рис. 9. Кадр с выделенными характеристическими точками.

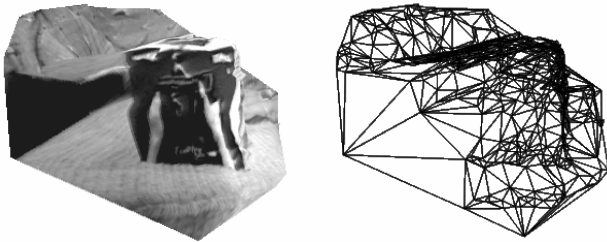


Рис. 10. Восстановленная VRML-модель с другой точки зрения, справа – то же в виде сетки.

8. ЗАКЛЮЧЕНИЕ

Разработана базовая составляющая системы восстановления трехмерных сцен по последовательности цифровых изображений. Путем численного моделирования найден критерий адаптивного выбора модели (приближения) вычислений. Проведены оценки точности восстановления. Предложен подход к выделению плотной подматрицы измерений, основанный на поиске клики графа. Показано, что применительно к аэрокосмическим системам достижима высокая точность восстановления рельефа, и точность восстановления ориентации камеры. Установлено, что наиболее слабым местом текущего состояния системы является требование малого изменения изображений между парами последовательных кадров, и планируемый блок грубого совмещения существенно необходим. В случае больших дальностей необходим альтернативный способ устранения неоднозначности формы

[3,4], основанный, например, на априорном грубом знании смещения камеры между некоторыми парами кадров.

9. ЛИТЕРАТУРА:

[1] *Carlo Tomasi, Takeo Kanade*. Shape and Motion from Image Streams: a Factorization Method, Part 3, Detection and Tracking of Point Features //Technical Report CMU-CS-91-132 / School of Computer Science, Carnegie Mellon University. — April 1991.

[2] *Conrad I. Poelman, Takeo Kanade*. A Paraperspective Factorization Method for Shape and Motion Recovery: //Technical Report CMU-CS-93-219 / School of Computer Science, Carnegie Mellon University. — 11 December 1993.

[3] *J. Costeria and T. Kanade*. A multi-body factorization method for motion analysis. Technical report, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213, September 1994. CMU-CS-TR-94-220.

[4] *Янова Н.В., Юрин Д.В.* Итеративный алгоритм восстановления трехмерных сцен, движения и фокусного расстояния камеры в перспективной проекции, основанный на факторизации матриц. //В сб. Труды конференции. 12-я Международная Конференция по Компьютерной Графике и Машинному Зрению ГрафиКон'2002 –С. 123–129. Нижний Новгород, 16-21 сентября 2002 г.

[5] *Shewchuk J.R.* Triangle: Engineering a 2d quality mesh generator and delaunay triangulator. First ACM Workshop on Applied Computational Geometry. 1996.

[6] The First EOS Satellite. NASA'S Earth Observing System EOS AM-1. http://terra.nasa.gov/Publications/AM-1_brochure.pdf

[7] *Stan Birchfield and Carlo Tomasi*. Multiway Cut for Stereo and Motion with Slanted Surfaces //Proceedings of the Seventh IEEE International Conference on Computer Vision, Kerkyra, Greece, September 1999.

[8] *Y. Xiong and S. Shafer*. Depth from focusing and defocusing. <http://www.cs.cmu.edu/yx/papers/IUW93.ps.Z>.

[9] *Vladimir Kolmogorov and Ramin Zabih*. Computing Visual Correspondence with Occlusions using Graph Cuts. // In: International Conference on Computer Vision, July 2001.

[10] *M. Watanabe S. K. Nayar and M. Noguchi*. Real-time focus range sensor. Technical report, Department of Computer Science, Columbia University, New York, USA, June 1994. CUCS-028-94.

[11] *C. Bron and J. Kerbosch*. Algorithm 457: Finding All Cliques of an Undirected Graph. CACM, 16(9):575-577, 1973.

Об авторах

Наталья Владимировна Свешникова – студентка 5-го курса ФФКЭ Московского Физико - Технического Института.
E-mail: oduduka@cos.ru

Андрей Сергеевич Чернышов – студент 4-го курса ФФКЭ Московского Физико - Технического Института.
E-mail: andrew@cos.ru

Юрин Дмитрий Владимирович, к.ф.-м.н., ФГУП НИИ «ОПТ-ЭКС», нач. отдела.
E-mail: yurin@cos.ru