

На правах рукописи

Конушин Антон Сергеевич

**Алгоритмы построения трехмерных компьютерных
моделей реальных объектов для систем
виртуальной реальности**

Специальность 05.13.11 – математическое и программное обеспечение
вычислительных машин, комплексов и компьютерных сетей

А в т о р е ф е р а т
диссертации на соискание ученой степени
кандидата физико-математических наук

Москва - 2005

Работа выполнена в Институте прикладной математики
им. М.В.Келдыша РАН

Научный руководитель - кандидат физико-математических наук, доцент
Баяковский Юрий Матвеевич

Официальные оппоненты:

- доктор физико-математических наук,
Кугушев Евгений Иванович
- кандидат физико-математических наук,
Иванов Денис Владимирович

Ведущая организация:

Государственный Научно–Исследовательский Институт Авиационных Систем
(ГосНИИАС)

Защита состоится « _____ » _____ 2005 г. в _____ часов
на заседании Диссертационного совета Д 002.024.01 в Институте
прикладной математики им. М.В. Келдыша РАН по адресу: 125047,
Москва, Миусская пл. 4.

С диссертацией можно ознакомиться в библиотеке Института прикладной
математики им. М.В. Келдыша РАН

Автореферат разослан « _____ » _____ 2005 г.

Ученый секретарь диссертационного совета
доктор физико-математических наук

Т.А. Полилова

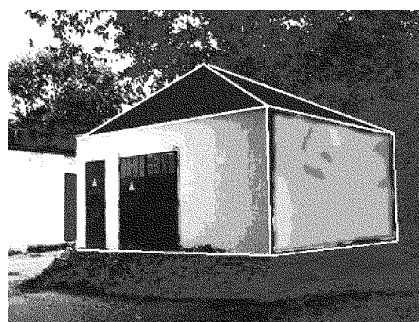
Общая характеристика работы

Актуальность работы

Термином «Виртуальная реальность» (VR) обозначают новые виды компьютерных интерфейсов, нацеленных на создание у пользователя ощущения реальности окружения. Это ощущение вынуждает пользователя взаимодействовать с окружающим миром точно так же, как он бы вел себя в реальности. Типичным примером систем VR являются тренажеры различных транспортных средств от автомобилей до пилотируемых космических аппаратов, достоверно воспроизводящие ощущения водителя или пилота, позволяющие обучать людей управлению ими, Рис. 1(а). Элементы виртуальной реальности также используются в интерфейсах, называемых «расширенной реальностью», когда в изображения реального мира добавляются синтезированные объекты. Последнее необходимо, например, при архитектурном планировании строительства и реконструкции зданий, Рис. 1(б).



(а)



(б)

Рис. 1 Интерфейсы типа "виртуальная реальность" (а) и "расширенная реальность" (б)

Достоверная визуализация виртуального мира играет важную роль при создании у пользователя ощущения реальности. Степень реалистичности изображения в настоящее время сильно зависит от точности и качества используемых моделей. В системах расширенной реальности для корректного встраивания виртуальных объектов в изображения необходимо точно определить положение камеры в пространстве, и оценить общее пространственное расположение наблюдаемых объектов.

Наиболее легкодоступным источником информации об объектах реального мира являются фотоизображения, поэтому в последние 15 лет большое внимание уделяется разработке алгоритмов и систем построения моделей реальных объектов по изображениям. Однако доведенные до коммерческого уровня системы, например, Canoma, ImageModeler, PhotoModeler, требуют

точного выделения вершин, ребер и границ объектов на фотоизображениях и сопоставления их вершинам, ребрам и границам выбранной модели простой формы. Этот процесс очень трудоемок, поэтому подобные системы не получили широкого распространения.

В сложившейся ситуации возникла очевидная потребность в автоматизированных системах построения трехмерных моделей по изображениям, не требующих дорогостоящей дополнительной аппаратуры, в которых взаимодействие с пользователем сводится к малому количеству простых операций.

Цель работы

Целью работы является исследование и разработка алгоритмов и методов построения трехмерных компьютерных моделей реальных объектов, позволяющих существенно снизить необходимый объем взаимодействия с пользователем по сравнению с существующими методами, а также разработка системы построения трехмерных моделей для апробации предложенных алгоритмов и методов.

Научная новизна работы

Предложен новый алгоритм робастной оценки параметров моделей, показано его преимущество по точности и устойчивости при решении задачи калибровки камеры по сравнению с существующими аналогичными методами. Разработан новый алгоритм отслеживания точечных особенностей, превосходящий по характеристикам существующие алгоритмы, особенно при применении к задаче вычисления траектории движения камеры. Предложен алгоритм построения следов, устойчивый к перепадам резкости в последовательности изображений. Разработаны новые алгоритмы определения параметров камеры и траектории ее движения.

Предложены новые методы построения моделей для объектов формы типа цилиндр и параллелепипед по набору зашумленных трехмерных точек, лежащих на их поверхности.

Практическая значимость и реализация

Разработаны и доведены до практической реализации методы и алгоритмы вычисления траектории движения фото- и видеокамеры и построения трехмерных объектов по последовательности изображений. Программные реализации описываемых в диссертации методов удовлетворяют всем требованиям и ограничениям, сформулированным при постановке задачи.

На основе разработанных автором алгоритмов разработана система построения моделей 3-х мерных объектов простой формы по фотографиям и видеопоследовательностям, не требующая большого объема взаимодействия с пользователем. Система разрабатывалась в Лаборатории Компьютерной Графики и Мультимедиа факультета ВМиК МГУ им. Ломоносова по заказу Samsung Advanced Institute of Technologies.

Апробация работы и публикации

Основные результаты работы докладывались и обсуждались на:

- 15-ой международной конференции по компьютерной графике и машинному зрению GraphiCon 2005, Россия, Новосибирск, 2005
- 12-ой международной конференции «Ломоносов-2005»
- 14-ой международной конференции по компьютерной графике и машинному зрению GraphiCon 2004, Россия, Москва, 2004
- 13-ой международной конференции по компьютерной графике и машинному зрению GraphiCon 2003, Россия, Москва, 2003
- Международной конференции по обработке изображений «International Conference on Image Processing 2002», США, Рочестер, 2002
- Семинаре по компьютерной графике и машинному зрению Ю.М.Баяковского (ф-т ВМиК МГУ)

Основные результаты работы изложены в 7 научных публикациях. Системы, в которые внедрены разработанные алгоритмы, защищены российскими и международными патентами.

Структура и объем работы

Диссертация состоит из введения, пяти глав, заключения, списка литературы и приложения. Содержание работы изложено на 158 страницах (включая 8 страниц приложения). Список литературы включает 91 наименование. В работе содержится 45 рисунков и 4 таблицы.

Содержание работы

В первой главе дается краткий обзор существующих систем реконструкции трехмерных моделей по набору изображений. В подобных системах модель конструируется пользователем из набора объектов простой формы. Параметры модели, такие как размеры и взаимное расположение блоков, определяются по

изображениям, для чего от пользователя требуется указать соответствие между вершинами и ребрами модели и их изображениями.

Разработанная система позволяет существенно снизить объем взаимодействия с пользователем благодаря автоматическому выделению точечных особенностей в изображениях, и вычислению соответствующих им трехмерных точек на поверхностях наблюдаемых объектов. При этом происходит вычисление траектории движения камеры, что позволяет использовать разработанную систему для разработки интерфейсов типа «расширенная реальность». Множество вычисленных трехмерных точек используется для подгонки параметрических моделей. Общая блок-схема работы предложенной системы приведена на Рис. 2



Рис. 2 Общая схема работы системы построения 3-х мерных моделей по набору изображений

Вторая глава посвящена робастным алгоритмам оценки параметров моделей, на которых основываются большинство алгоритмов, используемых для калибровки камеры и построения трехмерных моделей. В ней производится обзор существующих методов робастной оценки на основе случайных выборок и методов уточнения полученных оценок, а также описывается предложенный автором новый алгоритм решения данной задачи.

Одной из ключевых задач машинного зрения является сопоставление информации, содержащейся в изображениях, некоторой математической модели. Общую задачу оценки параметров модели по имеющимся данным можно записать следующим образом. Нужно определить вектор параметров θ ,

такой, что $F(X, \theta) = 0$; где $X = \{x_i\}, i = 1..N$ – набор точек исходных данных, $F(X, \theta)$ задает исследуемую математическую модель. Классическим примером задачи оценки параметров является аппроксимация набора точек прямой линией. В таком случае $X = \{x_i\}, i = 1..N$ – это набор точек, $F(x, \theta)$ – уравнение прямой, θ – параметры прямой. Исходные данные x *загрязнены*, т.е. помимо точек, порожденных исследуемой моделью, в данных содержатся *загрязняющие* точки, порожденные ошибками измерений, которые никак не зависят от исследуемой модели. Загрязняющие точки называются *выбросами* (outliers) Доля порожденных исследуемой моделью данных обозначается как $\gamma \in [0,1]$. Точки, порожденные исследуемой моделью зашумлены некоторой помехой ε , математическое ожидание которой $M(\varepsilon) = 0$, а дисперсия $D(\varepsilon) = \sigma^2, \sigma = const$. Метод оценки параметров модели по набору загрязненных исходных данных называется *робастным* методом оценки. Наибольшее распространение в области машинного зрения получили методы на основе случайных выборок. Первый подобный алгоритм RANdom SAMpling Consensus (RANSAC) был предложен в 1981 году Фишером и Болесом.

Для того, чтобы оценить параметры модели необходимо определить, какие точки из исходного набора являются загрязняющими выбросами, а какие нет. Однако в общем случае не существует способа определить заранее, является ли данная точка выбросом. Схема RANSAC решает эту проблему путем случайного перебора некоторого количества выборок из исходных данных. По каждой выборке оцениваются параметры модели θ . Предполагается, что значение функции $F(x, \theta)$ достигает своего минимума на гипотезе θ , оцененной по выборке, не содержащей выбросов. Каждая выборка строится случайным образом, т.е. из набора исходных данных последовательно выбирается нужное количество точек. Каждая из точек набора исходных данных берется с равной вероятностью. Для максимизации вероятности построения выборки без выбросов, размер выборки N берется минимально необходимым для оценки параметров модели.

Существующие алгоритмы робастной оценки параметров модели требуют заранее определенных параметров распределения шума в исходных данных. Во многих случаях эти параметры могут быть подобраны вручную или оценены используемым алгоритмом получения исходных данных. Однако зачастую это невозможно в связи с естественной флюктуацией значений параметров шума в исходных данных. Одним из примеров подобных задач является задача оценки положения и ориентации камеры в пространстве, называемых в совокупности *позой* камеры. Эта задача послужила стимулом для разработки нового общего

алгоритма робастной оценки параметров модели, адаптирующегося к уровню шума и доле загрязнения.

Для точной робастной оценки параметров моделей, в диссертационной работе предложен новый алгоритм AMLESAC (Adaptive Maximum Likelihood Estimation SAmping Consensus). Он основан на общей схеме робастной оценки параметров на основе случайных выборок, но дает оценку максимального правдоподобия гипотезы с одновременным вычислением параметров шума, с которым генерировались исходные данные. Алгоритм поиска вектора параметров модели θ с наибольшим правдоподобием опирается на предположение, что исходные данные X представляют собой смесь двух выборок. Первая выборка состоит из точек, удовлетворяющих исследуемой модели. Для таких точек ошибка согласования с моделью удовлетворяет нормальному распределению. Вторая выборка состоит из выбросов, ошибка согласования которых удовлетворяет равномерному распределению. Полная схема алгоритма может быть сформулирована следующим образом:

- 1) Вычислить количество итераций алгоритма M по априорной оценке γ .
- 2) Повторить M раз:
 - a. Построить случайную выборку исходных данных $S_k \subset X$.
 - b. Оценить гипотезу θ_k по выборке S_k .
 - c. Применить локальную оптимизацию к гипотезе θ_k .
 - d. Построить случайную выборку исходных данных $T_k \subset X$.
 - e. Оценить параметры смеси σ и γ для текущей гипотезы θ_k и выборки T_k .
 - f. Измерить правдоподобие гипотезы θ_k по всем исходным данным X с оцененными значениями σ и γ .
- 3) Выбрать гипотезу θ с максимальным правдоподобием.
- 4) С помощью метода нелинейной минимизации уточнить гипотезу θ

В качестве целевой функции используется логарифм правдоподобия исходных данных X для модели с параметрами θ :

$$-L = -\sum_i^N \log \left(\gamma \left(\frac{1}{\sqrt{2\pi}\sigma} \right)^s \exp \left(-\frac{e^2}{2\sigma^2} \right) + (1-\gamma) \frac{1}{v} \right)$$

где $e = \|\tilde{x} - x\|$ - это расстояние (норма L2) между измеренными координатами x и их математическим ожиданием \tilde{x} реальной точки \bar{x} . \tilde{x} - это такая точка, которая удовлетворяет уравнению $F(\tilde{x}, \theta) = 0$, и минимизирует L2 норму $e = \|\tilde{x} - x\|$, v - это объем пространства, внутри которого выбросы распределены равномерно.

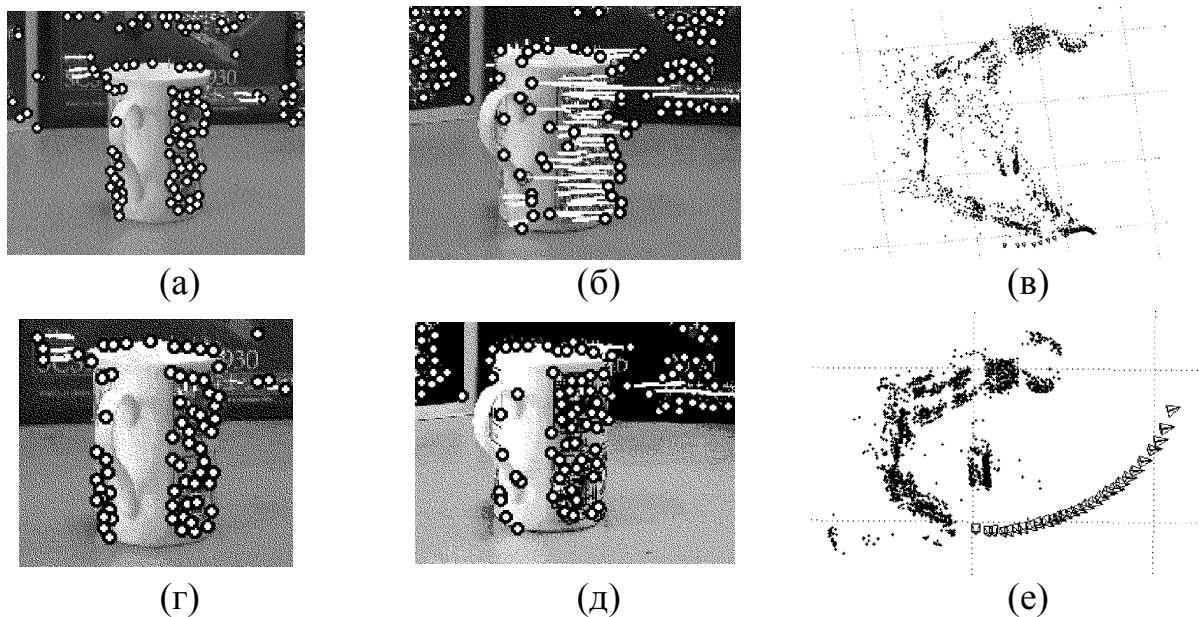


Рис. 3 Применение AMLESAC для задачи оценки позы камеры. (а),(б) Репроекция 3-х мерных точек на 6 и 9 кадры при оценке позы с помощью MSAC (г),(д) Репроекция 3-х мерных точек на 6 и 9 кадры при оценке позы с помощью AMLESAC (в).(е) Траектория движения камеры и 3-х мерные точки, вычисленные с помощью алгоритмов MSAC и AMLESAC соответственно. Алгоритм MSAC помечает точки цилиндра как выбросы, т.к. их ошибка репроекции оказывается большой, что ведет к существенным ошибкам в оценке траектории движения камеры

Значение правдоподобия гипотезы зависит от доли γ и дисперсии ошибок σ , поэтому точность оценки правдоподобия напрямую зависит от точности оценки этих параметров. Для каждого набора исходных данных значение этих параметров может быть разным, поэтому в каждом случае они должны оцениваться отдельно. В предлагаемом алгоритме для каждой гипотезы θ производится поиск в пространстве γ и σ таких значений, при которых правдоподобие θ достигает максимума. Для этого проверяются с некоторым шагом все варианты доли $\gamma_k \in [\gamma_{\min}, 1]$, начиная с минимального значения γ_{\min} . Для каждого γ_k оценка σ_k вычисляется с помощью медианы ошибок $\gamma_k * N$ точек с наименьшей ошибкой. По σ_k с помощью итеративного алгоритма вычисляется уточненное значение γ_k' . В качестве (γ, σ) выбирается пара (γ_k', σ_k) , соответствующая наименьшему значению логарифма правдоподобия, которая дополнительно уточняется методом градиентного спуска.

Для повышения точности алгоритма используется локальная оптимизация каждой гипотезы θ_k . Локальной оптимизацией называется оценка гипотезы на выборке s_{lo} , по размеру большей минимальной. Выборка s_{lo} строится из точек

x с наименьшей ошибкой $e = \|\tilde{x} - x\|$. Как показывают эксперименты, локальная оптимизация значительно повышает точность оценки параметров смеси (γ, σ) и самой гипотезы.

Для снижения вычислительной сложности алгоритма параметры смеси (γ, σ) оцениваются с использованием не всех исходных данных, а только части. Для этого случайным образом строится выборка $T_k \subset x$, и параметры оцениваются только по ней. Проведенные эксперименты показали, что размер выборки T_k может не превышать 100 точек, при любом размере набора исходных данных большем 100. Правдоподобие каждой гипотезы по-прежнему вычисляется на всем наборе исходных данных.

Количество итераций M выбирается таким образом, чтобы вероятность построения выборки без выбросов размером N с априорной оценкой максимальной доли выбросов в исходных данных $(1-\gamma)$ была больше заданной, обычно 95% - 97%.

Разработанный алгоритм был протестирован при решении ряда задач робастной оценки параметров, таких как подгонка прямых и определение позы камеры как на синтетических, так и на реальных исходных данных. Наибольшее преимущество перед существующими аналогами алгоритм продемонстрировал на реальных последовательностях изображений, полученных с фотокамеры Canon IXUS 500. Вследствие накопления ошибки, при построении траектории движения камеры начиная с некоторого кадра существующие алгоритмы давали ложный результат, приводящий к существенному отклонению вычисленной траектории от истинной, как показано на Рис. 3.

В третьей главе описываются методы отслеживания и построения следов точечных особенностей в последовательностях изображений, приводится обзор существующих алгоритмов и используемых в них методов.

Точечной особенностью x' изображения называется точка, чья окрестность отличается от окрестностей близлежащих точек по выбранной мере, т.е. $\{\forall x: |x'-x| < r \rightarrow \rho(\Omega_x, \Omega_{x'}) > \varepsilon\}$, где Ω_x - окрестность точки x , называемая *окном поиска*, а $\rho(\Omega_x, \Omega_{x'})$ - функция близости окрестностей по некоторой мере. Последовательность положений точечной особенности $\{x'_i\}, i = \overline{n_{start}, n_{end}}$, где i - номер кадра видеопоследовательности, называется *следом* точечной особенности. n_{start} - первый кадр, на котором обнаружена точечная особенность, n_{end} - последний кадр, на котором определено положение точечной

особенности, $n_{end} > n_{start}$. След точечной особенности $\{x'_i\}, i = \overline{n_{start}, n_{end}}$ называется корректным, если существует точка трехмерной сцены X , такая что $x'_i = P_i(X), i = \overline{n_{start}, n_{end}}$, где P_i - уравнение перспективной проекции, соответствующее i -ому изображению. Такая точка X называется точечной особенностью наблюдаемой сцены. Другими словами, корректный след точечной особенности представляет собой совокупность проекций некоторой точки сцены на сегмент кадров исходной последовательности. Выделение на последовательности изображений набора следов точечных особенностей устанавливает предполагаемое соответствие между изображениями и набором точечных особенностей сцены. *Соответствием* (match) называется пара точек x_1, x_2 на изображениях I_1, I_2 , которые соответствуют одной и той же точечной особенности x' . Фактически, любая пара точек x_{i_1}, x_{i_2} из следа точечной особенности $\{x'_i\}, i = \overline{n_{start}, n_{end}}$ является соответствием. Соответствие (x_{i_1}, x_{i_2}) называется корректным, если существует точка трехмерной сцены X , проекциями которой являются точки x_{i_1}, x_{i_2} .

Существует два метода поиска соответствий. Первым из них является алгоритм сопоставления особенностей, представляющий собой последовательный перебор всевозможных пар особенностей x_1, x_2 на изображениях I_1, I_2 , для поиска наиболее близких к друг другу по выбранной мере точечных особенностей. Второй метод называется отслеживанием особенностей. Для каждой точечной особенности x_1 изображения I_1 на втором изображении I_2 ищется точка x_2 , наиболее близкая к x_1 по выбранной мере. Для поиска точки x_2 используется итеративный алгоритм на основе метода градиентного спуска.

Для построения набора следов точечных особенностей предлагается новый алгоритм, совместно использующий оба метода поиска соответствий, что позволяет совместить достоинства обоих методов, и скомпенсировать ошибочно найденные ложные соответствия. Разработанный алгоритм построения набора следов разбивает всю исходную последовательность изображений ключевыми кадрами на сегменты, на каждом из которых изменяется либо только ориентация камеры в пространстве, (камера вращается вокруг своего оптического центра) либо меняется также положение камеры в пространстве. В каждом из этих двух случаев, движения точечных особенностей между кадрами внутри сегмента должно удовлетворять либо перспективному преобразованию плоскости, называемому гомографией, либо эпиполярному ограничению, записываемому в форме фундаментальной

матрицы. Для поиска нового ключевого кадра используется алгоритм отслеживания особенностей Лукаса-Канаде. Пусть I_i - последний ключевой кадр, тогда отслеживание продолжается до нового ключевого кадра I_j , такого, что оцененная на соответствиях точечных особенностей на паре кадров (I_i, I_j) фундаментальная матрица лучше согласуется с исходными данными, чем гомография. Для сравнения степени согласия используется информационный критерий GRIC. Если на протяжении 10 кадров фундаментальная матрица оказывается менее согласованной с построенным набором соответствий, то в качестве нового ключевого кадра выбирается кадр I_{i+10} . В этом случае корректным видом ограничения на этом сегменте признается гомография.

После определения нового ключевого кадра I_j производится сегментация ложных следов на сегменте $[I_i, I_j]$. Те из следов, которые не удовлетворяют корректному виду ограничений, признаются ложными и отбрасываются.

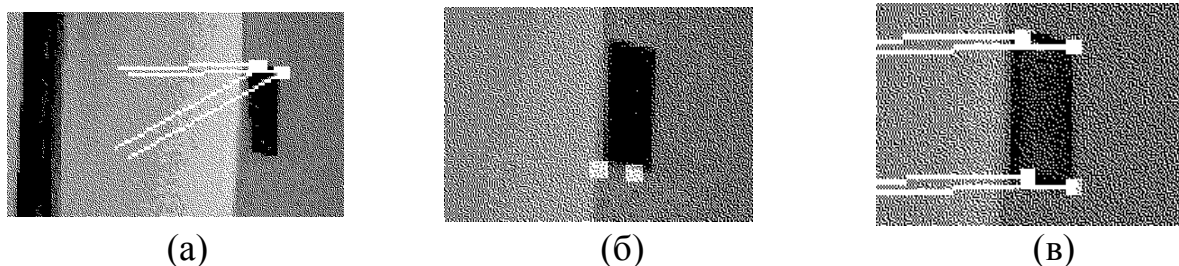


Рис. 4 Управляемое слежение. (а) Соответствия, найденные алгоритмом Лукаса-Канаде (б) Положение точек на новом кадре, предсказанное с помощью переноса гомографией (в) Результат работы управляемого слежения. Все углы рамки были корректно отслежены.

Одним из недостатков алгоритмов отслеживания точечных особенностей является постепенное изменение изображения окрестности точечной особенности во время слежения. Это приводит к накоплению ошибок измерения координат точки на текущем кадре и срыву слежения. Для уменьшения этой проблемы в разработанном алгоритме построения набора следов используется алгоритм отслеживания с коррекцией. Положение точечной особенности на новом кадре, определенное алгоритмом отслеживания Лукаса-Канаде, используется как предсказание положения точки на новом кадре. В малой окрестности предсказанного положения осуществляется поиск точечной особенности. Особенность, окно поиска которой наиболее близко к окну поиска особенности на предыдущем кадре, выбирается как скорректированное положение отслеживаемой точки на текущем кадре.

Алгоритмы отслеживания точечных особенностей в качестве начального приближения положения отслеживаемой точки на новом кадре используют ее положение на предыдущем кадре. В некоторых случаях, например, при наличии в сцене объектов с регулярным рисунком, на новом кадре рядом с предыдущим положением точки может оказаться другая точечная особенность, похожая по своей окрестности на отслеживаемую, но соответствующая другой 3-х мерной точке сцены. В этом случае алгоритм отслеживания выдаст эту ближайшую точку в качестве результата слежения, что приведет к срыву слежения и преждевременному завершению построения следа. Для решения этой проблемы начальное приближение должно задаваться более точно. В разработанном алгоритме такое приближение вычисляется с помощью переноса положения точки на предыдущем кадре гомографией, которая оценивается с помощью метода наименьших квадратов по всем корректным соответствиям точечных особенностей рассматриваемой пары кадров в некоторой окрестности данной точки. Этот алгоритм называется управляемым слежением. После определения нового ключевого кадра, управляемое слежение применяется для всех точечных особенностей, следы которых были отброшены как ложные на данном сегменте. Во многих случаях, когда это произошло по причине ошибки обычного алгоритма слежения, управляемое слежение позволяет корректно отследить точечную особенность на всем сегменте. Пример работы управляемого слежения приведен на Рис. 4.

В тех случаях, когда управляемое слежение не позволяет найти корректное соответствие, используется алгоритм управляемого сопоставления. Согласно последнему, проверяются всевозможные пары соответствий точечных особенностей на рассматриваемой паре кадров, которые удовлетворяют корректному виду ограничений (гомографии или фундаментальной матрице).

Если в сцене присутствуют объекты с высокодетализированной текстурой, то на их изображениях может быть найдено значительное количество точечных особенностей с высоким качеством. При использовании обычной схемы для инициализации следов выбираются точки с наибольшим качеством. В этом случае большинство следов инициализируется по точечным особенностям изображений этих объектов, что приводит к образованию доминантного подмножества, к некорректному определению вида ограничений на соответствия, и к отбрасыванию большого количества корректных следов как ложных. Для устранения этого недостатка в предложенном алгоритме используется адаптивная схема инициализации следов. Все изображение разбивается на прямоугольные области одинакового размера. Для

инициализации следов из каждой области выбирается одинаковое количество особенностей, вследствие чего возрастает равномерность распределения следов по изображению, снижается вероятность образования доминантного подмножества. Пример работы адаптивной схемы приведен на Рис. 5.



(а)



(б)

Рис. 5 Адаптивная схема инициализации следов. (а) Обычная схема инициализации следов. (б) Адаптивная схема инициализации следов.

Необходимо отметить, что на количество выбранных особенностей на салфетке значительно снизилось, а на коробке выросло. Соответственно следы более равномерно распределены по изображению.

Эксперименты на тестовых последовательностях изображений реальных сцен продемонстрировали превосходство разработанного алгоритма по средней длине следа точечной особенности и по равномерности распределения следов в изображении. Это преимущество позволило увеличить точность оценки траектории движения камеры. Количество же найденных следов несколько уменьшается, за счет снижения количества следов в высокодетализированных областях, а также за счет того, что в аналогичных алгоритмах при срыве слежения для той же самой точечной особенности след инициализируется повторно.

Существующие алгоритмы поиска соответствий не справляются с существенными перепадами резкости между кадрами последовательности, на участках с такими перепадами обычно происходит разрыв следов точечных особенностей. Для поиска соответствий на кадрах с перепадами резкости, а также для связывания следов точечных особенностей до и после дефектного участка, предлагается новый алгоритм. В этом алгоритме определяется степень резкости каждого кадра исходной последовательности. Для получения этой оценки к изображению применяется фильтр Собеля, каждая строка полученного изображения рассматривается как одномерная функция. В качестве оценки резкости используется среднее расстояние между экстремумами этой функции.

Как показывают эксперименты, количество найденных особенностей в изображении зависит от размера окна поиска. При этом существует зависимость между степенью резкости изображения и размером окна поиска. При поиске соответствий для выбранной пары кадров сравниваются оценки их резкости, и более резкое изображение размывается до уровня менее резкого. С учетом найденной зависимости выбирается оптимальный размер окна поиска, при котором количество найденных особенностей максимально. Производится поиск соответствий методом сопоставления среди наборов найденных точечных особенностей с выбранным окном поиска. Для сегментации ложных соответствий используется гомография, которая оценивается робастным алгоритмом.

По известным параметрам гомографии между каждой парой соседних кадров на дефектном участке можно оценить гомографию между первым и последним кадрами дефектного участка. Это ограничение используется для управляемого сопоставления точечных особенностей первого и последнего кадров. Найденные соответствия между конечными кадрами дефектного участка позволяют соединить следы точечных особенностей до и после дефектного участка.

Четвертая глава посвящена методам определения траектории движения камеры и построения множества 3-х мерных точек, лежащих на поверхностях наблюдаемых объектов. В качестве исходных данных используется набор следов точечных особенностей.

Траектория движения камеры обычно называется просто *движением* камеры, а множество 3-х мерных точек сцены – *структурой* сцены. Положение и ориентация камеры в пространстве называются *позой* камеры. Таким образом, задача совместного определения траектории камеры и множества 3-х мерных точек сцены называется задачей определения *структуры и движения*.

В работе реализованы последовательный и иерархический алгоритмы определения структуры и движения, и предложены усовершенствования используемых в них алгоритмов решения подзадач. Последовательный алгоритм основан на первоначальном выборе двух кадров, для камер которых вычисляется взаимная ориентация и положение в пространстве. По соответствиям точечных особенностей этой пары кадров инициализируется набор 3-х мерных точек структуры сцены. Затем последовательно вычисляются позы камер всех остальных кадров, при этом уточняются координаты точек структуры сцены, а также добавляются новые точки в структуру. Структура и движение затем уточняются с помощью метода связок. Иерархический

алгоритм разбивает исходную последовательность ключевыми кадрами на сегменты, на каждом из которых структура и движение вычисляются независимо последовательным алгоритмом. Затем траектории движения камеры на соседних сегментах объединяются, для чего вычисляется относительное положение и ориентация в пространстве камер последнего и первого кадров соседних сегментов. Процесс повторяется до тех пор, пока не будет вычислена траектория движения камеры на всей исходной последовательности изображений. Для повышения точности калибровки при уточнении методом связок на каждом сегменте используются только такие 3-мерные точки, следы которых определяются как корректные на протяжении всего сегмента.

Для оценки позы камеры используется подмножество точек структуры сцены, видимых с данного кадра, и набор их проекций, взятых из набора следов точечных особенностей. Для устойчивой к накоплению ошибок оценке позы камеры предлагается новый алгоритм оценки позы камеры, использующий алгоритм AMLESAC, автоматически адаптирующийся к уровню ошибок и загрязнения в исходных данных. Для оценки позы камеры по набору из 6 трехмерных точек и соответствующих им проекций без выбросов также предлагается новый линейный алгоритм. Он состоит из следующих шагов:

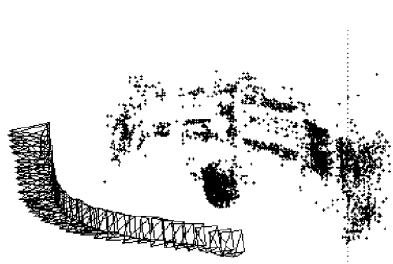
1. Вычисление проекций точек на идеальной плоскости камеры
2. Оценка матрицы перспективной проекции методом наименьших квадратов (МНК) путем составления и решения системы линейных уравнений
3. Вычисление ориентации камеры в пространстве по матрице перспективной проекции
4. Оценка положения камеры в пространстве методом МНК с учетом вычисленной ориентации камеры

После робастной оценки позы камеры, последняя уточняется с помощью метода нелинейной минимизации по всем точкам, которые были помечены робастным алгоритмом как удовлетворяющие модели. Однако каждая из 3-мерных точек, используемая при оценке позы камеры, может быть вычислена с различной точностью. Использование точек, содержащих большую ошибку, при уточнении позы может внести существенные погрешности в параметры позы. Для решения этой проблемы в качестве целевой функции предлагается использовать взвешенную сумму квадратов отклонений проекций трехмерных точек от измеренных значений. Вес каждой точки пропорционален оценке точности вычисления последней. Точность оценки координат 3-мерной точки

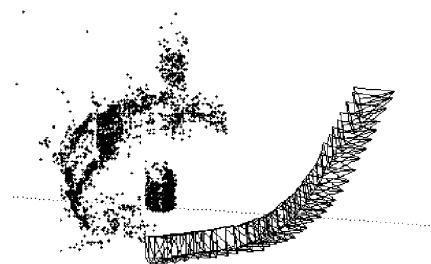
зависит от угла между лучами, исходящими из оптических центров камер, использованных при триангуляции, и проходящими через саму точку. Поэтому предложенный алгоритм оценивает вес каждой точки согласно углу между вышеописанными лучами.



(а)



(б)



(в)

Рис. 6 Пример вычисления траектории движения камеры и структуры сцены предложенным алгоритмом. (а) Одно из исходных изображений (б),(в)

Результат работы алгоритма

Проведенная серия экспериментов на синтетических и реальных данных продемонстрировала превосходство предложенных алгоритмов над существующими аналогами по устойчивости и точности оценки. Пример работы алгоритма приведен на Рис. 6.

В пятой главе описываются два подхода к построению трехмерной модели наблюдаемых объектов: на основе методов подгонки примитивов и с использованием методов плотного стерео, а также приводит описание некоторых алгоритмов обработки построенных моделей.

Первый подход опирается на параметрическое представление поверхности объектов простой формы. Подгонкой параметрической модели называется поиск таких значений параметров, при которых модель наиболее согласована с имеющимися исходными данными. В работе предлагается два алгоритма для подгонки параметрических моделей цилиндра и параллелепипеда. В качестве исходных данных используются трехмерные точки структуры, вычисленные описанным в четвертой главе алгоритмом. Процесс построения модели объекта на основе параметрической модели состоит из следующих шагов:

- Выделение пользователем изображения моделируемого объекта на исходных изображениях
- Выбор пользователем типа параметрической модели
- Поиск начального приближения значений параметров модели
- Уточнение значений параметров

Интересующий пользователя объект выделяется на нескольких исходных изображениях с помощью прямоугольной рамки, при этом из всего множества трехмерных точек выделяется подмножество, состоящее из всех точек, принадлежащих объекту, и некоторой доле не принадлежащих, представляющих собой загрязняющую помеху.

Предложенный алгоритм подгонки модели цилиндра в качестве параметров использует положение оси цилиндра в пространстве, высоту и радиус цилиндра. На первом шаге находится начальное приближение оси цилиндра. Пусть $\{p_i\}, i = \overline{1, M}$ - набор точек на поверхности тела вращения, $\{n_i\}, i = \overline{1, M}$ - нормали к поверхности тела вращения. Тогда ось вращения тела вращения пересекает большинство линий вида $\{p_i + \alpha n_i\}$. Оценка множества нормалей производится путем аппроксимации плоскостью множества точек из окрестности выбранной точки. Аппроксимация производится с помощью робастного алгоритма MSAC, причем нормаль оценивается только по таким плоскостям, степень согласования которых с множеством точек из окрестности не превышает заданного порога. По построенным нормальям строится множество прямых вида $\{p_i + \alpha n_i\}$, используемых как исходные данные для оценки оси цилиндра. Оценка производится робастным алгоритмом MSAC. Каждая гипотеза выбирается как прямая, сумма расстояний от которой до всех прямых из выборки минимальная. На втором шаге верхняя и нижняя границы цилиндра определяются как верхняя и нижняя границы множества проекций точек цилиндра на ось. Радиус цилиндра вычисляется как среднее расстояние от оси цилиндра до точек, не лежащих на его верхней и нижней границах.

После того, как получено начальное приближение значений параметров цилиндра, они уточняются с помощью нелинейного алгоритма минимизации Гаусса-Ньютона или Левенберга-Макардта. Пусть Cyl - поверхность цилиндра. В качестве целевой функции используется взвешенная сумма расстояний от точек до поверхности цилиндра:

$$S = \sum w(\rho(Cyl, p_i)), \quad w(x) = \begin{cases} x, & x < \Delta \\ \Delta, & x \geq \Delta \end{cases}$$

где $w(\cdot)$ - робастная весовая функция, называемая М-оценкой, предназначенная для минимизации влияния загрязняющей помехи, т.е. таких точек, которые не лежат на поверхности цилиндра.

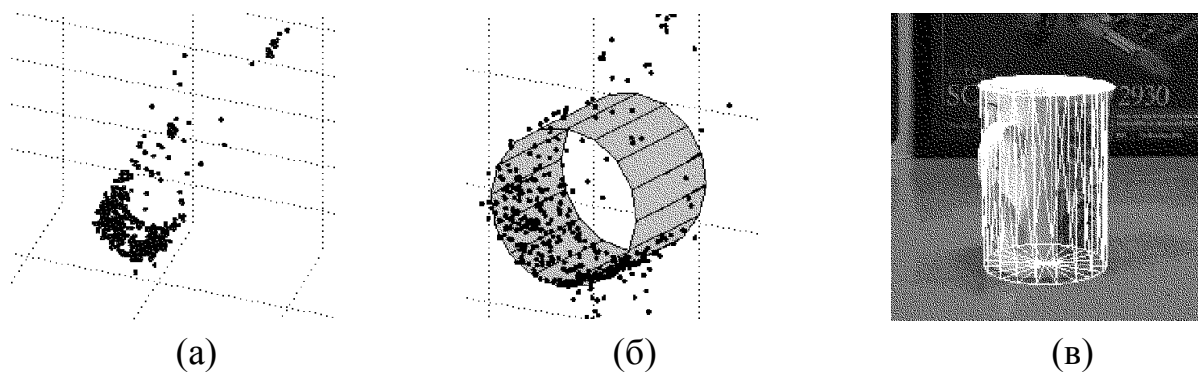


Рис. 7 Подгонка цилиндра. (а) Исходное множество точек (б) Построенная параметрическая модель (в) Проекция сеточной модели на исходное изображение

Рис. 7 иллюстрирует процесс подгонки модели цилиндра к набору трехмерных точек, полученных по последовательности изображений реальной сцены.

Общую схему разработанного алгоритма для подгонки параметрических моделей объектов в форме параллелепипеда можно записать так:

- Выделение плоскостей видимых граней с помощью робастного алгоритма
- Определение осей параллелепипеда
- Определение границ граней параллелепипеда
- Определение невидимых граней по вычисленным границам видимых

Плоскости в наборе точек выделяются с помощью итеративного применения робастного алгоритма MSAC, каждая плоскость-гипотеза проводится по $k > 3$ точкам с помощью метода МНК. Оси параллелепипеда определяются путем вычисления множества нормалей к видимым граням и их векторного произведения. Границы грани оцениваются как границы множества проекций точек, принадлежащих грани, на оси параллелепипеда. На Рис. 8 приведен пример работы алгоритма на наборе 3-х мерных точек, полученных с последовательности изображений реальной сцены.

Если моделируемый объект должен иметь форму прямоугольного параллелепипеда, тогда после выделения первой плоскости следующая плоскость проводится по выборкам из 2-х точек таким образом, чтобы ее нормаль была перпендикулярна нормали первой плоскости. Нормаль к третьей плоскости до ее выделения можно вычислить как векторное произведение нормалей первых двух плоскостей.

Для вычисления текстуры поверхности моделируемого объекта по полученной параметрической модели строится сеточная модель поверхности объекта.

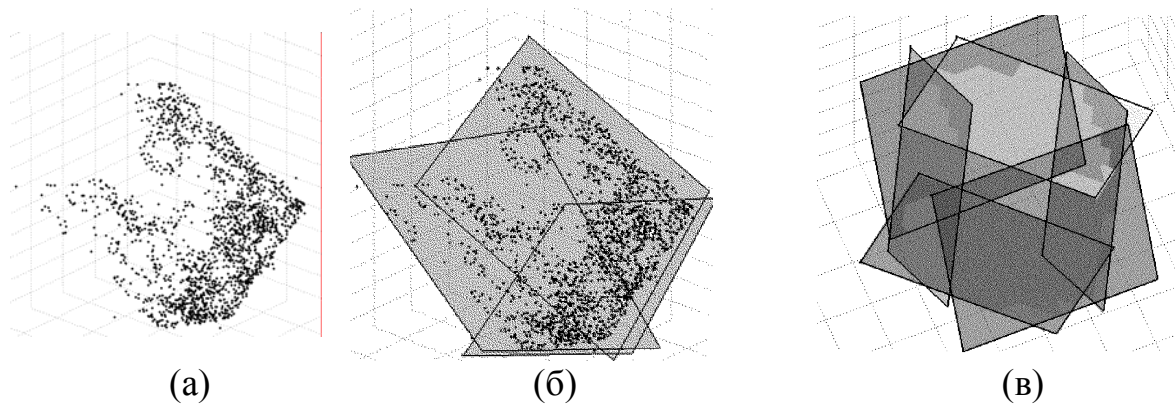


Рис. 8 Инициализация параметров параллелепипеда (а) Исходное множество точек (б) Найденные робастным алгоритмом плоскости (в) Все плоскости параллелепипеда

Второй подход к реконструкции трехмерных моделей использует представление моделей в виде набора изображений с картами глубины. Подобные модели называются *основанными на изображениях* (image-based models). Совместно с другими сотрудниками Лаборатории Компьютерной Графики МГУ автором был разработан формат записи таких моделей, который был принят в стандарт AFX MPEG4. Разработанный формат опирается на иерархическое представление сцены, используемое в языке описания трехмерных моделей и сцен VRML, и совместим с бинарным форматом сцен BIFS, являющимся основой стандарта MPEG4. Алгоритм построения таких моделей формулируется следующим образом:

- Вычисление карт глубины для всех исходных изображений $\{I_i\}, i = \overline{1, N}$ с помощью алгоритма плотного стерео
- Совместная обработка набора изображений с картами глубины для получения модели объекта

Плотным стерео называется задача поиска для каждого пикселя на одном изображении I_i соответствующего ему пикселя на другом изображении I_j . После установления попиксельного соответствия между изображениями, для каждого пикселя вычисляется положение соответствующей ему 3-х мерной точки с помощью алгоритма триангуляции.

Большинство точек сцены видно сразу на нескольких изображениях исходной последовательности, поэтому одной и той же точке сцены может соответствовать сразу несколько трехмерных точек с разных изображений с глубиной. Эта дублирующаяся информация значительно увеличивает размер модели из набора изображений с глубиной, снижает скорость ее визуализации,

а также четкость визуализации. Для удаления избыточной информации разработан алгоритм, основанный на вычислении качества каждой трехмерной точки и удалении из каждой группы близкорасположенных точек всех, кроме одной с наибольшим качеством. После удаления этих точек, цвет соответствующих им пикселей изображений меняется таким образом, чтобы повысить качество изображений при сжатии с потерями, например, по алгоритму стандарта JPEG. Предложенный алгоритм позволяет эффективно сжимать модели в виде набора изображений с глубиной, и повысить скорость визуализации.

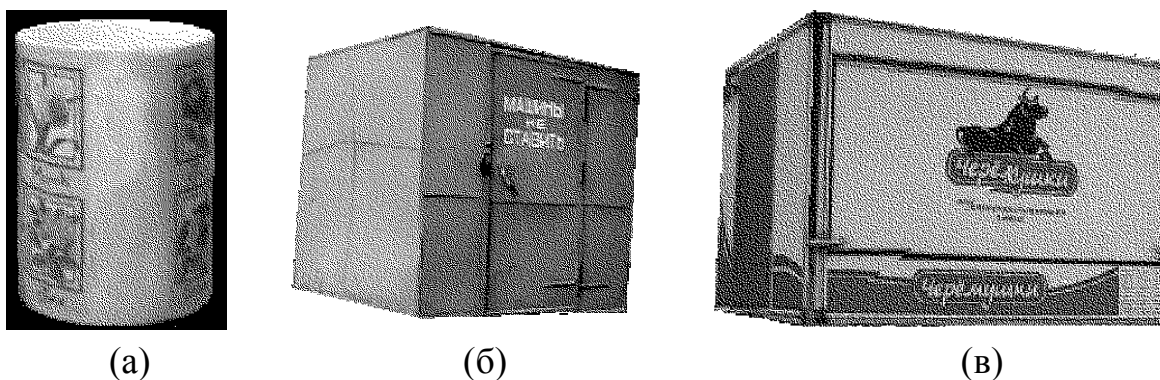


Рис. 9 Примеры реконструированных моделей объектов реального мира

Для моделей на основе изображений с картами глубины также предложен алгоритм редактирования текстуры путем модифицирования визуализированных изображений объектов, что для многих пользователей гораздо проще непосредственного редактирования текстуры. Разработанный алгоритм основан на обратном распространении изменений от модифицированного изображения ко всем изображениям с глубиной. Пользователь модифицирует одно изображение, после чего алгоритм определяет во всех исходных изображениях области, соответствующие части поверхности объекта с модифицированной текстурой. Цвет каждого пикселя из данных областей изменяется таким образом, чтобы соответствовать модифицированной текстуре.

На Рис. 9 приведены примеры построенных моделей реальных объектов.

В заключении сформулированы основные результаты работы.

В приложении описываются модели и свойства перспективной проекции и ограничения, накладываемые на взаимное расположение проекций трехмерной точки на двух изображениях перспективной проекцией при условии статичности сцены. На эти свойства и их модели опираются все методы определения траектории движения камеры и трехмерной реконструкции.

Основные результаты диссертационной работы:

1. Предложены новые алгоритмы построения следов точечных особенностей на последовательности фотоизображений или видеопоследовательности, которые позволяют добиться более высокой точности определения траектории движения камеры.
2. По результатам проведенного исследования предложен новый робастный алгоритм оценки параметров моделей на основе случайных выборок. Алгоритм адаптируется к уровню шума и позволяет добиться высокой точности и устойчивости в условиях значительного уровня шума и ложных данных.
3. Разработан комплекс алгоритмов для генерации трехмерной модели объектов простой формы типа цилиндр и параллелепипед по облаку трехмерных точек. Алгоритмы успешно функционируют на загрязненных данных, что позволяет сократить объем взаимодействия с пользователем для уточнения полученных моделей.
4. На основе предложенных алгоритмов разработана программная система построения трехмерных компьютерных моделей реальных объектов по набору фотоизображений или видеопоследовательности для формирования виртуальной среды.

Публикации по теме диссертации

1. A.Konouchine, V.Gaganov, V.Vezhnevets "AMLESAC: A New Maximum Likelihood Robust Estimator", Proc. Graphicon-2005, pp, 93-100, Novosibirsk, June 2005
2. A.Konouchine, V.Gaganov, V.Vezhnevets "Combined Guided Tracking and Matching with Adaptive Track Initialization", Proc. Graphicon-2005, pp. 301-304, Novosibirsk, June 2005
3. А.Конущин «Система построения трехмерных моделей реальных объектов по последовательности изображений» (тезисы), с.30, Москва, Ломоносов-2005
4. А.Конущин, К.Мариничев, В.Вежневек "Обзор робастных схем оценки параметров моделей на основе случайных выборок". Труды конференции Graphicon-2004, с.275-278, Москва, сентябрь 2004.
5. Е.Лисицин, А.Конущин, В.Вежневек "Отслеживание точечных особенностей в видеопоследовательностях с изменениями резкости". Труды конференции Graphicon-2004, с.233-236, Москва, сентябрь 2004.
6. L. Levkovich-Maslyuk, A.Ignatenko, A.Zhirkov, A.Konushin, I.K.Park, M. Han, Y. Bayakovski. "Depth Image-Based Representation and Compression for Static and Animated 3-D Objects". IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 14, NO. 7, pp. 1032-1045, JULY 2004.
7. A.Ignatenko, A.Konushin. "A Framework for Depth Image-Based Modeling and Rendering". Proc. Graphicon-2003, pp. 169-172, Moscow, Russia, September 2003.
8. Y. Bayakovski, L. Levkovich-Maslyuk, A. Ignatenko, A. Konushin, D. Timasov, A. Zhirkov, M.Han, I.K.Park "Depth Image-Based representations for static and animated 3D objects". IEEE 2002 International Conference on Image Processing, vol.3, pp.25-28, Rochester, New York, September 22-25, 2002