

СИНТЕЗ ИЗОБРАЖЕНИЙ ТРЕХМЕРНЫХ ОБЪЕКТОВ СЦЕН НА ОСНОВЕ ИХ ФОТОГРАФИЙ

Зори С.А.

Донецкий Национальный Технический Университет
Кафедра Прикладной Математики и Информатики
zori@pmi.dgtu.donetsk.ua

В роботі дано огляд існуючих алгоритмів синтезу зображень складних тривимірних об'єктів за допомогою фотографічних зображень. Сформульовано основні недоліки та особливості цих методів та можливі області їхнього застосування.

Введение

Задача создания изображений трехмерных объектов, с максимальной точностью повторяющих их реальные прототипы (синтез с высокой степенью реализма), возникает довольно часто. Понятие «системы рендеринга на основе изображений» (image-based rendering system) становится все более популярным как в сообществе компьютерной графики, так и в сообществе компьютерного зрения (computer vision community) [1-2]. Общая идея состоит в том, чтобы избежать дорогого в вычислительном отношении полного расчета 3D-модели сложного (в идеале - реального) объекта, а вместо этого использовать ряд связанных изображений объекта (сцены) как представление, из которого новые виды могут синтезироваться путем интерполяции (деформации) изображения - «морфинга».

Целью исследований является определение способности генерации изображений сложных объектов, информация о трехмерной модели которого не задана, а представлена только совокупностью типовых изображений объекта с различных ракурсов. Предполагается рассмотрение проблемы при известном (заданном) количестве изображений и исследование вопросов разработки эффективного алгоритмического обеспечения с целью выполнения процесса синтеза за заданное время, а также получения, по возможности, приближенной трехмерной модели объекта.

1 Основные методы синтеза на основе изображений

Общий принцип синтеза на основе изображений проиллюстрирован следующим рисунком (в примере - 2 исходных изображения).

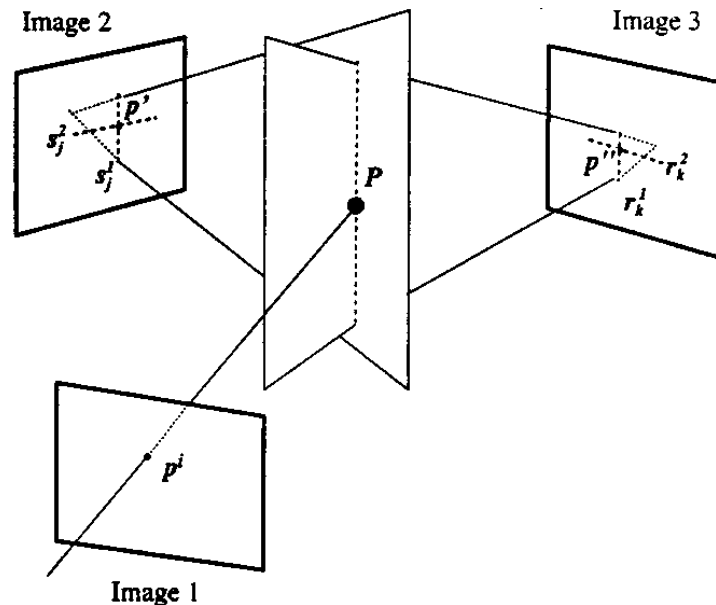


Рисунок 1 – Общая постановка задачи синтеза на основе изображений

Анализ литературных источников показывает, что работа в обозначенной области идет, в основном, по трем направлениям:

- интерполяция изображений;
- автономный (мозаично-основанный) синтез;
- интерактивный синтез.

1.1 Интерполяция изображений

Первый класс методов - интерполяция изображений – объединяет алгоритмы, создающие "промежуточные" изображения из двух или более исходных изображений. Сюда можно отнести морфинг изображений (morphing) [6, 7], прямую интерполяцию потоков изображения ("многомерный морфинг") [8-10], интерполяцию изображений, используя 3-D модели вместо потока изображения [2], и "физически корректную" интерполяцию изображений [6]. Все, кроме последних двух подгрупп, не гарантируют, что произвели физически правильные изображения, и ни один из методов не позволяет проводить экстраполяцию набора исходных данных – то есть создавать виды, которые находятся вне конуса видимости исходных изображений.

Seitz и Dyer [6, 10] показали, что можно выполнять интерполяцию по опорной линии (base-line) пары изображений и получать физически

правильные изображения (в отличие от интерполяции на основе потоков - flow-based). Их подход предполагает вначале анализ и выравнивание изображений, интерполирование по опорным линиям (которые являются параллельными после исправления), и конечный рендеринг. К сожалению, только изображения вдоль линии, соединяющей два образцовых изображения, могут быть сгенерированы таким образом, что не позволяет пользователю свободно «виртуально двигаться» в пространстве.

В отличие от интерполяции на основе поля потоков (flow-field) исходных изображений, можно проводить интерполяцию с использованием функции светового потока (plenoptic) [7, 10], которая представляет собой количество света, испускаемого каждой точкой в пространстве как функцию от направления. В работах Levoy и Hanrahan, Gortler и другие [10] интерполируют между плотным множеством из нескольких тысяч (!) исходных изображений, чтобы восстановить plenoptic- функцию. То есть, происходит значительное увеличение числа требуемых изображений, для вычисления оптических потоков между исходными изображениями.

1.2 Автономный синтез изображений

Во втором классе, автономном синтезе на основе мозаики (mosaic-based), синтез осуществляется следующим образом - берется множество накладывающихся изображений сцены и "сшиваются" вместе. Самое простое сшивание происходит, когда движение камеры включает только вращение - в этом случае, преобразование между видами является параметрическим и не включает никакую трехмерную форму (преобразование, являющееся 2D- проективным преобразованием, homography). Примером является хорошо известная технология "QuickTime VR" [2, 3]. Szeliski и Kang [7] создают мозаики с высоким разрешением из видео-потоков с низкой разрешающей способностью, Peleg и German [6] используют только фиксированную камеру, представляя множество проекций изображений с нее. Недостаток этого класса состоит в том, что нельзя корректно моделировать свободное перемещение камеры с набором исходных изображений.

1.3 Интерактивная интерполяция и синтез

Главным ограничением вышеупомянутых методов является то, что для их использования требуется относительно большое количество изображений, представляющих объект. Третий класс, интерактивный синтез, уменьшает число необходимых изображений, используя «3D-из-2D геометрию» [2], чтобы получить некоторую интерактивную функцию деформации из набора исходных изображений для создания новых видов «на лету» в соответствии с пользовательскими установками параметров

позиции виртуальной камеры. Laveau и Faugeras [6] были первыми, кто использовал такой подход для синтеза новых видов, это позволило им даже проводить экстраполяцию видов.

Использование этого подхода предполагает наложение дополнительных ограничений (Eipolar-ограничений) к процессу синтеза. Эти ограничения обусловлены особенностям, которые возникают при некоторых движениях камеры (например, когда виртуальный центр камеры является коллинеарным с центрами камер исходных видов), и требуется дополнительное определение точек соответствия (matching points) на объекте и их проекций на исходных изображениях (с помощью eipolar-линий). Эти ограничения могут быть компенсированы использованием карты глубины среды. McMillan и Bishop [6] используют полную карту глубины (3D реконструкция движения камеры и среды) вместе с eipolar-ограничением, чтобы обеспечить прямое соответствие между виртуальным движением камеры и движением по проекциям. Карты глубины легко обеспечиваются для синтетических сред, в то время как, для реальных сцен - процесс неустойчив (нет информации о глубине) [6, 7].

Требования, предъявляемые к "оптимальным" интерактивным методам синтеза, изложены ниже:

- Неявное моделирование сцены - чтобы уменьшить, в максимальной степени, число вычислительных шагов от момента получения исходного соответствия начальных изображений до получения необходимых алгебраических структур, которые удовлетворяли бы требованиям синтеза новых видов. Поэтому желательно, чтобы параметры камеры оставались в максимальной степени неявными в процессе;

- Несингулярные конфигурации - использовать функции деформации, которые являются свободными от особенностей движения камеры.

- Режим запуска - параметры виртуальной позиции камеры должны быть интуитивно просты для пользователя. Например, вращение и перемещение камеры от ее текущей позиции наиболее распространены среди средств 3D просмотра существующих интерактивных синтезирующих систем.

Ни один из существующих подходов для интерактивного синтеза не удовлетворяет всем трем требованиям. Например, можно удовлетворять первому требованию за счет усложнения режима запуска, определяя контрольные точки; использование карт глубины обеспечивает интуитивный режим запуска и отсутствие сингулярностей, но не удовлетворяет требованию неявного моделирования сцены.

В последнее время появился ряд работ, названный методами тензоров [8, 9, 10]. Это подходы, основанные на соединении трилинейных

деформирующих функций, которые оставляют сцену и параметры камеры неявными. В это же время подход не зависит от особенностей движения камеры и управляется широко распространенным режимом запуска, используемым большинством современных средств просмотра 3D сцен.

Синтез нового вида основан на следующей парадигме: третий вид может быть сгенерирован однозначно на основе двух исходных изображений и полученного тензора (матрицы соответствия) при помощи деформирующей функции. Деформирующая функция управляет изменениями в коэффициентах тензора в результате перемещения виртуальной камеры.

Тензор (trilinear tensor) – это матрица $3 \times 3 \times 3$. Необходимо по крайней мере семь точек соответствия среди трех изображений, чтобы линейно восстановить трилинейный тензор. Восстановленный один раз тензор может использоваться для репроекции, поэтому, имея два изображения и тензор, третье изображение может быть уникально определено и синтезировано посредством деформирующей функции, приложенной к двум исходным изображениям. Таким образом, необходимо выполнение следующих шагов:

- установить точное соответствие между парой изображений;
- восстановить базовый тензор (трилинейного тензора, в случае трех исходных изображений, или тензорного расширения фундаментальной матрицы, в случае двух изображений или каскадного тензора в случае большего количества изображений);
- реализовать корректный механизм репроекции;
- реализовать обработку возможных коллизий при синтезе нового вида.

Метод тензоров свободен от описанных выше ограничений, обладает высокой точностью и, поэтому, признается наиболее перспективным.

Таким образом, анализ литературных источников позволяет провести следующую приблизительную классификацию методов синтеза на основе фотографий (рис. 2).



Рисунок 2 – Классификация методов синтеза на основе изображений

2 Основные области применения синтеза на основе фотографий

Впервые заявив о себе миру в клипе "Black & White" М. Джексона, компьютерный морфинг прочно обосновался в арсенале выразительных средств, используемых сегодня в самых разнообразных областях применения средств компьютерной индустрии - в кинематографии, рекламных роликах, мультипликации, компьютерных играх, средствах виртуальной реальности.

В результате проведенных исследований можно определить некоторые основные области применения методов синтеза на основе изображений:

- Видеоморфинг – при создании и компоновке видеосюжетов;
- Медицинская томография – при недостаточно подробной последовательности томографических сечений;
- Распознавание и реконструкция изображений;
- Видеокомпрессия – для сжатия оцифрованной видеoinформации;

- Метеорология – получение более точного прогноза погоды по последовательности спутниковых фотографий атмосферы;
- Геодезия – получение изображений рельефов и ландшафтов (а, также, в идеале, их трехмерных моделей) по набору спутниковых фотографий;
- Тренажеры транспортных средств – реалистичное моделирование в реальном времени окружающей обстановки в ситуациях, когда цена ошибки оператора транспортного средства недопустимо высока;
- Игры и системы виртуальной реальности – синтез в реальном времени новых изображений по существующим с целью увеличения степени интерактивности и реалистичности.

Доступный набор средств, предоставляемых современными компьютерными программами для морфирования изображений, примерно одинаков и требует от аниматоров кропотливого и рутинного труда по заданию соответствий между аналогичными областями начального и конечного изображений ("якорными точками"). Попиксельная интерполяция между заданными «якорными точками» векторами позволяет построить векторное поле, используемое для трансформации изображений при просчете промежуточных фаз морфирования. Среди доступных специализированных приложений встречаются предназначенные только для обработки поверхности объектов или, наоборот, создания одного типа объектов [11]. Особняком стоит группа программ, выполняющих вспомогательную роль. Это, например, Scanoma или Photomodeler. Их основное предназначение — создавать трехмерные объекты из фотографий, генерировать фотореалистичные текстуры, т.е. делать заготовки для дальнейшего использования в других приложениях.

К сожалению, в литературе, как правило, отсутствует обзор и описание инструментальных средств, использующихся при синтезе на основе изображений в рассмотренных выше областях, а тем более методов, лежащих в их основе.

Заключение

В данной работе дан обзор и выполнена классификация основных методов синтеза на основе фотографий, показаны их основные особенности и недостатки, обозначены основные области применения методов. Дана характеристика основных инструментальных средств, используемых для генерации новых изображений по фотографиям. В дальнейшем предполагается выполнить разработку и моделирование эффективного алгоритмического обеспечения с целью выполнения процесса качественного синтеза на основе изображений за заданное время, и исследование его характеристик.

Литература

1. Сидоренко Ю. Три измерения трёхмерной графики. Компьютерное обозрение, 1999, №34, с 38-40.
2. Сидоренко Ю. Третье измерение фотографии. Компьютерное обозрение, 1999, №27, с.28-29.
3. Потапов М. 3D-мир по Фаренгейту. Компьютерное обозрение, 1999, №44, с 30-33.
4. Theo Pavlidis. Algorithms for graphics and image processing. Computer science press, 1985.-400p.
5. Kaufman, A. (Ed.): Volume Visualization. IEEE Computer Society Press, Los Alamitos, CA, 1999.
6. S. Carlsson. Duality of reconstruction and positioning from projective views. In Proceedings of the workshop on Science Representations, Cambridge, MA, June 1995.
7. Robust Recovery of Camera Rotation from three Frames. B. Rousso, A. Sashua, S. Peleg. Computer Graphics, 22, 4 (1998), 61-72
8. S. Avidan, A. Shashua. Novel view synthesis in tensor space/ CS report, CIS-9602, Technion, 1996
9. Shashua A., Avidan S. Novel view synthesis by cascading trilinear tensors. IEEE transactionson visualization and computer graphics, vol. 4, no. 4, pp. 293-306.
10. www.ntu.edu.sg/home/assourin/computer_graphics/tensors.html
11. Сидоренко Ю. Carrara: моделер, аниматор и не только. Компьютерное обозрение, 2000, №8, с 26-29.

Поступила в редакцию 12.01.04