

Метод автоматического синтеза панорамных изображений

Л. В. Серебряная, email: l_silver@mail.ru¹

А. И. Простаков, email: andrewprostakov@gmail.com²

^{1,2} Белорусский государственный университет информатики и радиоэлектроники

Аннотация. Работа посвящена синтезу панорам из отдельных изображений. Для этого применяется технология компьютерного зрения. Анализ и обработка изображений выполняются с использованием особых точек. Изложены особенности этапов построения панорам, а также математические модели и методы, положенные в их основу. Предложены алгоритмы регистрации сегментов изображений и сравнения сегментов. Создано программное средство, предназначенное для автоматического синтеза панорам.

Ключевые слова: компьютерное зрение, панорама, изображение, сегмент, особые точки, дескриптор, склейка изображений.

Введение

Работа посвящена построению панорамных изображений. Это представление пространства с большими углом и глубиной обзора. Панорамы удобно использовать для наглядного отображения пространств со множеством деталей. Существует особый вид панорам, позволяющих создавать виртуальные трехмерные сцены, что может найти применение в различных областях и сферах жизни.

Большинство панорам имеют вид фотографий. Объективы фотоаппаратов или камер ограничены в охвате углов обзора, вследствие чего они неспособны отразить всю требуемую картину. Поэтому при съемке создают набор фотографий с различными углами поворота камеры, а затем фотографии обрабатывают и объединяют в панораму.

Процесс пост-обработки изображений может выполняться разными способами, которые делятся на ручные и автоматизированные. Каждый вид методов имеет как достоинства, так и недостатки. Ручной способ оказывается более медленным и трудоемким, но дает качественный результат. Автоматизированный – быстрый, но при нем может пострадать качество. Поэтому необходим баланс между ручными и автоматизированными действиями в ходе построения панорам.

Для автоматизации построения панорам удобно использовать технологию компьютерного зрения [1]. Она позволяет получать

информацию из изображений и обрабатывать ее. Обработка изображений связана с построением трехмерных сцен на основе двумерных объектов.

В алгоритмах построения трехмерной структуры пространства используется понятие «особые точки» или «характерные точки». Это точки изображения, обладающие высокой локальной информативностью. Обычно выделение характерных точек является начальным этапом работы любой системы компьютерного зрения. Основным преимуществом использования особых точек считается относительная простота и скорость их обнаружения. К характерным точкам предъявляют ряд требований в виде их свойств: инвариантности, стабильности, уникальности, интерпретируемости.

Обзор программных средств (ПС) Photoshop, Image Composite Editor, PTGui, Autopano Pro и Autopano Giga, предназначенных для построения панорамных изображений, позволил оценить их преимущества и недостатки, а также сделать соответствующие выводы. Все рассмотренные приложения имеют схожие принципы использования и требования к исходным изображениям, такие как:

- доля перекрытий изображений должна быть приблизительно 20-60%;
- фото выполняются из одной точки расположения камеры методом поворота вокруг фиксированной оси;
- фото имеют одинаковый уровень яркости.

Следовательно, во всех ПС применяются схожие методы объединения и обработки изображений. Значит актуальной задачей является поиск методов построения панорам, которые позволят полностью или частично устранить вышеперечисленные ограничения.

1. Модели и методы, положенные в основу синтеза панорам

Фактически панорамное изображение синтезируется в результате объединения нескольких исходных изображений (сегментов). Процесс построения панорам состоит из трех этапов: регистрации, калибровки, сочетания. Каждый из них имеет специфические проблемы и методы их решения [2].

Регистрация подразумевает анализ исходных сегментов панорамы. Этот этап пропускается при использовании метода прямой подгонки. Калибровка – соотнесение сегментов друг с другом. Сочетание – объединение или склейка сегментов в одно изображение и ликвидация признаков выполненной процедуры.

Одной из фундаментальных операций при склейке изображений является их пространственное преобразование. Исходные сегменты представляют собой проекции некоторой сцены на плоскость объектива

в разных позициях, углах наклона и поворота. Сегменты не хранят подобную информацию, однако её наличие является необходимым условием для воссоздания сцены в виде панорамы. В этом контексте задачу объединения можно выразить как нахождение такого преобразования для каждого сегмента, которое позволит корректно разместить его на панораме.

Обработка изображений выполняется на основе их локальных особенностей, которые связаны с особыми или характерными точками. Необходимым свойством особенности является стойкость к преобразованиям.

В основу алгоритма поиска особых точек положен метод SIFT [3]. Это алгоритм для детектирования и описания локальных особенностей изображения. SIFT способен точно идентифицировать объекты в сложных ситуациях, например, при перекрытии. Deskриптор SIFT полностью инвариантен к вращению и масштабированию, частично – к аффинным преобразованиям и смене освещения. Особые точки определяются путем построения пирамиды масштабов для каждого изображения и нахождения локальных экстремумов интенсивности их пикселей. Затем точки последовательно уточняются. В результате deskриптор имеет вид вектора. После описания особенностей каждого сегмента выполняется процедура их сравнения, имеющая целью нахождение совпадений и вычисление соответствующих им преобразований.

Рассмотрим математическую модель, относящуюся к перечисленным действиям. Сначала строится пирамида масштабов изображения при помощи фильтра Гаусса. Это широко применяемый фильтр сглаживания, определяемый формулой

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}, \quad (1)$$

где x и y – координаты текущего пикселя относительно сглаживаемого (ядра размытия); σ – коэффициент стандартного отклонения.

Из значений функции $G(x,y)$ образуется матрица свертки, налагаемая на окрестность изменяемого пикселя:

$$I(x, y) = \sum_{dx, dy} I(x + dx, y + dy) * G(dx, dy). \quad (2)$$

Для обеспечения максимальной точности матрица свертки должна покрывать всё изображение, однако, ввиду специфики функции Гаусса,

вклад наиболее удаленных пикселей практически равен нулю по сравнению со вкладом ближайших, поэтому величину матрицы обычно ограничивают длиной стороны, равной 6σ .

Для нахождения особых точек на основе исходного изображения строится пирамида гауссианов, а затем – пирамида разностей гауссианов (DoG, Difference of Gaussian). Подобная операция позволяет локализовать градиенты интенсивности по их модулям или разбивать изображение на частотные составляющие. Одним из применений разностей гауссианов является выделение граней на картинке. Также они позволяют определять особые точки по их частотным свойствам. Таким образом, построенная пирамида DoG служит для поиска локальных экстремумов значения разности. Не все из найденных точек пригодны в роли особых. Для отбора каждая из них пропускается через ряд проверок. Предварительно следует уточнить координаты каждой точки, так как при масштабировании они получают некоторое отклонение. Для этого применяется интерполяция функции DoG в виде квадратичного многочлена Тейлора с ядром в точке экстремума. Если значение функции интерполяции $D(x)$ в точке x меньше 0,03, точка характеризуется низкой величиной контраста и отбрасывается.

Следующая проверка ликвидирует точки на гранях объектов. Подобные точки являются ярко выраженными экстремумами функции DoG, однако они весьма чувствительны к шумам. Граничные точки характеризуются тем, что взаимно перпендикулярные градиенты в них значительно различаются по величине. Это определяется при помощи матрицы Гессе, с помощью которой решается, будет ли отбракован рассматриваемый экстремум.

Оставшиеся уточненные экстремумы считаются особыми точками.

Дескриптор особой точки обеспечивает её стойкость к различным преобразованиям, например, к вращению и масштабированию. Для построения дескриптора необходима операция по вычислению компонентов градиента произвольной точки:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2}, \quad (3)$$

$$\theta(x, y) = \operatorname{tg}^{-1} \frac{L(x+1, y) - L(x-1, y)}{L(x, y+1) - L(x, y-1)}, \quad (4)$$

где $L(x, y)$ – значение гауссиана с фиксированным значением σ ; m – величина градиента; θ – его направление.

Сначала определяется окрестность точки, образующая дескриптор. Затем строится дескриптор, представляющий собой массив из 128 чисел.

После того, как получен набор особых точек для исходных сегментов, необходимо соотнести их между собой оптимальным образом.

Высока вероятность нахождения случайных пар ключевых точек, которые приведут к конфликту при вычислении трансформации. Во избежание таких ситуаций используется алгоритм RANSAC [4]. Результат работы RANSAC – достоверный оператор преобразования координат, определяющий положение одного сегмента панорамы относительно другого. Параметры преобразования могут быть уточнены.

При склейке изображений часто образуются области перекрытия, не совпадающие на 100%. Чтобы шов был незаметен, стараются так подобрать сегменты с изображениями, чтобы разность между пикселями смежных сегментов была минимальной. В случае появления небольшого шума он легко ликвидируется с помощью медианного пространственного фильтра. Еще одна задача – это отображение панорамы на плоскость изображения. При проецировании сферического изображения на плоскую поверхность неизбежны искажения. При малых углах обзора они будут практически незаметны, а при больших углах искажения значительны, что может вызывать неудобство восприятия изображений. Решением данной проблемы является применение различных видов проецирования на двухмерную плоскость [5].

2. Алгоритмы построения панорам

Регистрация сегментов. В ходе проведенных теоретических исследований были проанализированы алгоритмы, позволяющие сравнивать несколько изображений и объединять их в панораму. При этом остались еще не решенные вопросы. Один из них – каким образом связи между парами сегментов преобразуются в связи сегментов с общей панорамой. Наиболее часто для ответа на данный вопрос применяются два метода: аккумулирующая регистрация и регистрация с помощью распределения общей ошибки.

Под ошибкой здесь понимается разница между корректным и вычисленным расположением сегмента в контексте всей панорамы. Первый из методов имеет существенный недостаток, заключающийся в последовательном накоплении ошибки, которое в итоге может сильно исказить результат. Второй метод абстрактен и имеет недостаточное

количество примеров конкретных реализаций. Следовательно, необходима доработка и приспособление вышеуказанных методов.

В работе использован метод глобальной регистрации сегментов, основанный на аккумулирующем подходе. Его отличие состоит в том, что связи между изображениями представляют собой не линейную цепь, а структуру данных типа «дерево». Регистрация сегментов осуществляется рекурсивно, начиная с корня дерева, и далее по всем его потомкам. Таким образом, ошибка при вычислении расположения каждого сегмента зависит не от всех вышерасположенных элементов, а только от стоящих на пути от корня дерева к нему. Это позволяет строить дерево, используя те связи между сегментами, которые характеризуются наименьшим значением ошибки. На практике ошибка напрямую зависит от схожести сегментов, т.е. величины их перекрытия и особенностей области перекрытия. Чем более схожи изображения, тем точнее определяются между ними связи.

Генерация панорамы возможна сразу после получения исходных изображений. Необходимой информацией для генерации являются оценки схожести сегментов. Когда все отношения определены, начинается процесс создания дерева. Ошибка при выводе каждого сегмента зависит от суммарной ошибки его связи с корнем дерева. Следовательно, целесообразно выбрать корнем тот сегмент, который характеризуется наибольшей степенью схожести с другими сегментами.

Далее осуществляется переход к построению второго уровня дерева. Для этого среди всего множества еще не зарегистрированных в панораме сегментов выбираются те, для которых корень является наиболее схожим сегментом. Они напрямую связываются с корнем и для каждого вычисляется ошибка. Все последующие сегменты будут зависеть от ошибок этих связей. Для каждого полученного элемента второго уровня иерархии выполняется та же операция, что и для корня – поиск и присоединение ближайших сегментов. Процесс построения дерева продолжается до тех пор, пока имеются незарегистрированные сегменты для добавления в иерархию.

Сравнение сегментов панорамы. Для работы процедуры генерации панорамы необходимыми данными являются оценки схожести всех исходных изображений. От фактической схожести сегментов зависит точность их объединения в панораме. Следовательно, процедуры сравнения двух сегментов и определения трансформации между ними должны быть связаны между собой.

Как уже было определено, каждый сегмент панорамы задается конечным множеством особых точек, которые находятся с помощью метода SIFT. По этим точкам определяются локальные особенности

каждого изображения. Если сегменты имеют схожие особенности, то делается вывод о присутствии на них одного или схожих объектов, и, следовательно, о схожести самих сегментов.

В ходе попарного сравнения всех сегментов для каждой точки одного из них находится наиболее схожая с ней точка на другом изображении методом поиска ближайшего соседа. Все пары «соседей» запоминаются и сортируются по величине расстояния между собой. Метрика расстояния между точками выбрана таким образом, чтобы пары, описывающие один объект на двух изображениях, имели минимальное расстояние. Тогда среди упорядоченного множества пар можно выделить те, которые с наибольшей вероятностью представляют общую часть двух изображений. Все результаты метода SIFT сохраняются для дальнейшего использования в процедуре генерации панорамы.

Для преобразования одной группы точек в другую и объединения сегментов использован метод RANSAC.

3. Программное средство построения панорам

Разработанные алгоритмы были положены в основу создания ПС построения панорам на основе отдельных изображений.

Существуют общепринятые рекомендации по подбору изображений для их объединения в панораму. Все изображения должны быть связаны между собой перекрывающимися областями. Если используемые изображения – фотографии, то они должны иметь как можно более близкие значения баланса белого цвета. В идеальном случае – это фотографии объектов, сделанные из одной точки, так минимизируются искажения.

На рис. 1 приведены примеры изображений, из которых будет строиться панорама.



Рис. 1. Отдельные изображения для построения панорамы

На рис. 2 показана синтезированная панорама. Качественная панорама сохраняется, а дефектная – исправляется. Для этого

выбираются сегменты и исправляются связи между ними, поскольку изображения склеиваются посредством связей.

Если на результирующем изображении будут присутствовать значительные дефекты от склейки сегментов панорамы, то можно исправить результат. Чаще всего дефекты являются следствием неверного порядка следования сегментов или их неточной склейки. Некачественное изображение можно удалить из панорамы, после чего перезапустить процесс ее генерации из других сегментов.



Рис. 2. Синтезированное изображение

Заключение

Были изучены технологии компьютерного зрения и обработки изображений. Проанализированы прикладные ПС, предназначенные для решения задач, схожих с задачами данной работы, а также используемые в них модели и методы. В результате для построения панорам были выбраны наиболее подходящие модели и алгоритмы.

Создано программное средство, получающее в качестве исходных данных изображения и синтезирующее из них панорамы. Эффективные математические методы и алгоритмы позволяют использовать изображения с различными соотношениями сторон, углами наклона и масштабами. Величина областей перекрытия, разница в цветовой гамме и в искажении определяют качество созданной панорамы, однако их можно корректировать с помощью программного средства.

Результаты работы, полученные в ходе ее выполнения, показывают, что все задачи решены и цель достигнута.

Список литературы

1. Основы стереозрения [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/post/130300/>

2. Wikipedia [Электронный ресурс]. – Режим доступа:
http://en.wikipedia.org/wiki/Image_stitching
3. Habrahabr [Электронный ресурс]. – Режим доступа:
<http://habrahabr.ru/post/106302/>
4. Wikipedia [Электронный ресурс]. – Режим доступа:
https://en.wikipedia.org/wiki/Random_sample_consensus
5. Cambridge in Colour [Электронный ресурс]. – Режим доступа:
<http://www.cambridgeincolour.com/tutorials/image-projections.htm>