# Тема проекта

Здравствуйте уважаемая комиссия. Я Пронин А.С. из группы ИУ7-52б, и тема моей научно-следовательской работы “Обзор методов построения объемного   
изображения по стереопаре”.

# Введение

Видео и фото тесно вплелись в нашу жизнь. Почти каждый мобильный телефон оснащен камерой. Почти каждая камера умеет записывать видео. Повсеместно распространилась 3D-графика. С развитием возможностей усиливается потребность в “дешевом” построении 3D-сцен. Самый очевидный из таких методов – получение трехмерной картины мира по видеоряду или нескольким изображениям.

Есть и другие возможные применения. В киноиндустрии – для создания спецэффектов и 3D-фильмов, в военном деле – для измерения расстояний и т.д.

# Цель

Цель работы – анализ методов, применяемых при построении объёмного изображения по стереопаре. А поставленные задачи вы видите на слайде.

# Анализ предметной области

Как уже упоминалось выше, одна из основных задач стереозрения – это получение объёмного изображения по нескольким изображениям.

Данную задачу можно разбить на 3 этапа:

* поиск соответствующих точек на изображениях;
* получение трехмерных координат точек;
* построение трехмерной модели.

# Базовые понятия

В ходе работы, мною были рассмотрены основные термины данной предметной области, а именно:

1. **Проективная геометрия и однородные координаты**

* точки проективной плоскости;

на проективной плоскости точки описываются трехкомпонентным вектором

* прямые на проективной плоскости;
* трехмерное проективное пространство;
* проективное преобразование.

1. **Модель проективной камеры**

Проективная камера определяется центром камеры, главной осью — лучом начинающимся в центре камеры и направленным туда, куда камера смотрит, плоскостью изображения — плоскостью на которую выполняется проецирование точек, и системой координат на этой плоскости.

1. **Эпиполярная геометрия**

Эпиполярную геометрию используют для поиска стереопар, и для проверки того, что пара точек может быть стереопарой (т.е. проекцией некоторой точки пространства).

1. **Триангуляция точек**

Процесс определения трехмерных координат точки по координатам ее проекций в литературе называется триангуляцией (triangulation).

1. **Карта глубины**

Карта глубины (depth map) — это изображение, на котором для каждого пикселя, вместо цвета, храниться его расстояние до камеры.

# Существующие решения

Кроме этого, я провёл обзор следующих алгоритмов решающие задачи стереозрения:

* Алгоритм RANSAC
* Алгоритм SIFT
* Алгоритм SURF
* Метод поиска соответствующих точек, использующий DSI

# Алгоритм RANSAC

Алгоритм RANdom SAmple Consensus (RANSAC), предложенный Фишлером и Боллес, представляет собой стабильный метод оценки параметров модели на основе случайных выборок устойчевый к зашумлённости исходных данных (большая доля выбросов).

Часто возникает задача обработки данных, в которой необходимо определить параметры модели, которая должна удовлетворять исходным данным. Все исходные данные можно разделить на два типа: хорошие точки, удовлетворяющие модели, «не-выбросы» или «инлаеры» (англ. inlier) и ложные точки, шумы — случайные включения в исходные данные, «выбросы» или «аутлаеры» (англ. outlier).

Весь алгоритм состоит из одного цикла, каждую итерацию которого можно логически разделить на два этапа, которые представлены на слайде.

Преимуществом алгоритма RANSAC является его способность дать надёжную оценку параметров модели, то есть возможность оценить параметры модели с высокой точностью, даже если в исходном наборе данных присутствует значительное количество выбросов.

Одним из недостатков метода RANSAC является отсутствие верхней границы времени, необходимого для вычисления параметров модели. Если использовать в качестве некоторой границы времени максимальное число итераций, полученное решение может быть не оптимальным, а также существует очень малая вероятность, что ни одна модель не будет соответствовать исходным данным.

Алгоритм RANSAC часто используется в компьютерном зрении, например, для решения задачи сопоставления изображений и оценки фундаментальной матрицы для определения параметров расположения камеры.

# Алгоритм SIFT

Масштабно-инвариантная трансформация признаков – это алгоритм компьютерного зрения, используемый для выявления и описания локальных признаков в изображениях. Этот алгоритм был опубликован Дэвидом Лоу в 1999 году.

Сначала в SIFT извлекаются ключевые точки объектов из набора контрольных изображений и запоминаются в базе данных. Объект распознаётся в новом изображении путём сравнивания каждого признака из нового изображения с признаками из базы данных и нахождения признаков-кандидатов на основе евклидова расстояния между векторами признаков. Из полного набора соответствий в новом изображении отбираются поднаборы ключевых точек, которые наиболее хорошо согласуются с объектом по его местоположению, масштабу и ориентации. Определение подходящих блоков признаков осуществляется быстро с помощью эффективной реализации хеш-таблицы обобщённого преобразования Хафа. Каждый блок из 3 или более признаков, согласующийся с объектом и его положением, подлежит дальнейшей подробной проверке соответствия модели, и резко отклоняющиеся блоки отбрасываются. Наконец, вычисляется вероятность, что определённый набор признаков говорит о присутствии объекта, что даёт информацию о точности совпадения и числе возможных промахов. Объекты, которые проходят все эти тесты, могут считаться правильными с высокой степенью уверенности.

На слайде вы можете видеть на какие шаги, Лоу разбил свой алгоритм.

Данный алгоритм можно применять для следующих целей:

* распознавание объектов;
* определение местоположения и отслеживание техники/роботов;
* объединение изображений;
* моделирование 3D-сцен, распознавание и трассировка;
* распознавания действий человека;
* анализ человеческого мозга в трёхмерных изображениях Магнитно-резонансной томографии.

# Алгоритм SURF

Алгоритм SURF является улучшением алгоритма SIFT. Стандартная версия SURF в несколько раз быстрее, чем SIFT.

Основным результатом алгоритма SURF являются обнаруженные ключевые точки, которые должны иметь следующие свойства: повторяемость (для обеспечения возможности поиска такой же точки на следующем кадре); инвариантность к повороту и масштабу; уникальность.

Алгоритм работы SURF предусматривает выполнение этапов, представленных на слайде.

Обнаружение особых точек алгоритмом SURF основано на вычислении детерминанта матрицы Гессе (гессиана H). Значение гессиана используется для нахождения локального минимума или максимума яркости изображения.

Данный алгоритм также можно использовать для таких задач, как распознавание объектов, регистрация изображений, классификация или 3D-реконструкция.

# Метод поиска соответствующих точек, использующий DSI

DSI (disparity space image – несоответствие пространственного изображения)

Рассмотрим подробнее одномерный метод поиска соответствующих точек, использующий DSI. Пусть 𝑠𝑖 и 𝑠’𝑖 – соответствующие 𝑖-е строки левого и правого изображений, 𝐼(𝑥,𝑦) и 𝐼’(𝑥,𝑦) – соответствующие функции интенсивности. Тогда параметр 𝐷𝑆𝐼𝑖 вычисляется как разность интенсивностей:

𝐷𝑆𝐼𝑖(𝑥,𝑑) = ||𝐼(𝑥,𝑖) − 𝐼′(𝑥 − 𝑑,𝑖)||,

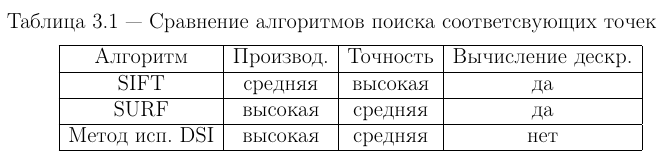
Другим способом построения 𝐷𝑆𝐼𝑖 является вычисление величины, основанной на корреляции функций интенсивностей вокруг пикселей (𝑥,𝑖) и (𝑥 − 𝑑,𝑖).

Задача сводится к поиску оптимального пути на полученной двумерной матрице. При этом за каждый тип движения назначается определенный штраф.

Задача решается методом динамического программирования. Если на изображении заранее известны GCP-точки (ground control points – точки, положение и соответствие которых мы можем определить достаточно точно), то количество возможных путей сокращается за счет использования следующего ограничения. Путь должен проходить так, чтобы он не противоречил расстановке базовых GCP-точек. Возможны также многозначные GCP-точки, когда точке одного изображения может соответствовать одна точка из небольшого набора вариантов на другом изображении.

# Сравнительный анализ алгоритмов

На таблице на слайде представлено сравнение рассмотренных методов поиска соответствующих точек по производительности в реальном времени, точности локализации ключевых точек и по наличию шагу вычисления дескриптора.



По сравнению с методом, использующим DSI, алгоритмы SIFT и SURF используют базы данных, и поэтому могут найти объект на новом изображении, в то время, когда первому методу всегда требуется 2 изображения. Кроме этого, SIFT и SURF позволяют не только представить трёхмерную модель объекта, но и распознать его.

Если же сравнивать алгоритмы SIFT и SURF, то наиболее предпочтительным оказывается SURF, т.к. за небольшой объём вычислений, с более высокой скоростью получаются почти те же ключевые точки, как у метода SIFT

Алгоритм RANSAC исключен из сравнения, т.к. он решает другую задачу. Алгоритм RANSAC можно использовать в совокупности с одним из вышеперечисленных для фильтрации полученных сопоставлений, потому что иногда сопоставляемые точки могут быть соотнесены некорректно, а RANSAC отсеивает «выпадающие» (outliers) из статистики точки.

# Заключение

По итогу проделанной работы была достигнута цель – проведён анализ методов, применяемых при построении объёмного изображения по стереопаре.

Также были решены все поставленные задачи, перечисленные на экране.