

Санкт-Петербургский государственный университет
Кафедра компьютерных технологий и систем

Авдюшкина Людмила Сергеевна

Магистерская диссертация

Построение 3D модели по набору изображений

Направление подготовки: 02.04.02

"Фундаментальная информатика и информационные технологии"

Образовательная программа: "Автоматизация научных исследований"

Научный руководитель,
кандидат физ.-мат. наук,
доцент
Погожев С.В.

Санкт-Петербург
2018

Содержание

Введение	3
Постановка задачи	5
Обзор литературы	8
Глава 1. Реконструкция на основе одного изображения	13
Глава 2. Особые точки (interested points)	17
2.1. Детекторы особых точек	18
2.1.1. ORB	18
2.1.2. BRISK	20
2.1.3. SIFT	23
2.1.4. SURF	25
2.2. Сопоставление особых точек	27
2.2.1. Метрики	27
2.2.2. Алгоритмы сопоставления	28
2.2.3. Фильтрация ложных соответствий	29
2.3. Реализация	30
Глава 3. Калибровка камеры	32
Глава 4. Объединение моделей	34
Глава 5. Представление трехмерных моделей	38
Глава 6. Программная реализация	40
Выводы	42
Заключение	43
Список литературы	45

Введение

Из всех способов получения и передачи информации визуальный способ – один из самых древних и широко применяемых. Из наскальных рисунков и декоративной керамики, служивших не только для украшения, но и для передачи информации, изображения постепенно превратились в источник конструктивных знаний. Примером этого могут служить астрономические карты и схемы, поразительно подробные схемы систем органов человеческого тела авторства Леонардо Да Винчи, а также схемы его изобретений. Изображения воспринимаются человеческим мозгом гораздо быстрее текстовой информации. Кроме того, схемы, графики, планы позволяют отбросить ненужную информацию и отразить только интересующие свойства объекта. Именно поэтому этот способ представления информации так часто применяется в учебниках и научных работах.

Но мы живем в трехмерном мире, а это значит, что при перенесении реального объекта на изображение, часть информации обязательно теряется. Например, архитекторам, чтобы отразить всю информацию о здании, приходится чертить планы в разных проекциях. А человеку, чтобы получить полную информацию, приходится выполнять обратные действия, а именно, мысленно собирать объект из его проекций. Таким образом, двумерные изображения хороши для отражения частного, но их недостаточно для получения общей картины объекта. В этом случае им на смену приходят трёхмерные модели.

На сегодняшний день трёхмерные модели применяются в самых разных областях. Например, архитекторы, помимо конструктивных схем, представляют еще и трехмерную модель, которая может быть как в виде реального макета, так и в виде компьютерной модели. Художники используют

упрощенные трехмерные модели для обучения. Хирурги используют точные трехмерные модели для планирования операций, а также для тренировки.

Одной из возможных областей применения трехмерных моделей является медицина, а точнее травматология. Статистика травматизма утверждает, что переломы конечностей – одни из самых распространенных типов переломов среди людей всех возрастов. Ранее для иммобилизации поврежденных конечностей накладывались гипсовые повязки. Однако они были тяжелыми, неудобными в носке, требовали предотвращения попадания на них воды, наложение таких повязок тоже было весьма дискомфортным, кроме того, гипс плохо пропускает рентгеновские лучи. На данный момент все более популярными становятся пластиковые лангеты, лишенные всех вышеперечисленных недостатков. С распространением 3D принтеров стала возможной печать фиксатора по индивидуальным параметрам непосредственно в больнице или травмпункте. Для этого необходима возможность быстрого получения трехмерной модели поврежденной конечности. В данной работе рассматривается частный случай задачи получения 3D модели по набору изображений на примере построения модели предплечья с запястьем по фото с разных ракурсов.

Постановка задачи

Определение 1. *Трехмерной реконструкцией* называют процесс получения координат каждой точки объекта или сцены в пространстве R^3 .

По отношению к объекту реконструкции методы получения трехмерной модели можно разделить на:

- **Активные методы**, взаимодействующие непосредственно с объектом реконструкции, будь то отдельный предмет или сцена полностью. Оборудование, осуществляющее реконструкцию активным методом, находится в пределах сцены. Примерами применения активной реконструкции можно назвать использование ультразвуковых, микроволновых и лазерных датчиков. Лазерные датчики лежат в основе достаточно широко используемых лазерных 3D-сканеров. Существенным недостатком применения активных методов является высокая стоимость используемого оборудования, его недостаточная мобильность и невозможность его использования для получения представления динамических сцен. Однако с помощью активной реконструкции возможно получение точной трехмерной модели, чего нельзя добиться применением пассивных методов.
- **Пассивные методы** не оказывают влияния на объект реконструкции. Вместо этого используется набор изображений или видеопоток, на основе которых восстанавливается трехмерная сцена или объект путем применения соответствующих алгоритмов. В отличие от активных методов, пассивные не могут гарантировать точность полученной модели. Однако они дешевы, мобильны, не требуют специального оборудования и навыков обращения с ними. Кроме того, пассивные методы пригодны для восстановления динамических сцен, что существенно расширяет

диапазон их возможного применения.

Очевидно, что во всех случаях, когда не требуется высокоточная реконструкция объекта или сцены, предпочтительнее использование пассивных методов реконструкции. Кроме того, применение методов активной реконструкции может быть ограничено ввиду особых свойств объекта реконструкции. Например, применение механического дальномера неприемлемо для упругих или легко деформируемых предметов. А на хрупких музейных экспонатах может негативно отразиться ультразвук.

В данной работе будет рассмотрена трехмерная реконструкция предплечья с целью дальнейшего использования полученной модели для трехмерной печати пластиковой лангеты. Очевидно, что применение для данной задачи методов активной реконструкции, а именно 3D сканеров, сопряжено с целым рядом сложностей:

- Высокая цена сканирующего оборудования
- Низкая скорость сканирования
- Необходимость наличия навыков обращения со сканирующим оборудованием

Таким образом, оптимальным решением будет использование методов пассивной реконструкции. Данный подход более удобен для пользователей, так как не требует специального оборудования (все современные смартфоны оснащены камерами), а следовательно, и специальных навыков обращения с ним. Кроме того, процесс фотографирования поврежденной конечности не доставит возможному пациенту дополнительных неудобств.

В данной работе рассматривается задача трёхмерной реконструкции предплечья с помощью методов пассивной реконструкции, а именно путём ана-

лиза набора фото с разных ракурсов. Особенностью задачи является отсутствие выраженных углов, и как следствие этого, сложность при выявлении особых точек и их сопоставлении. Данные особенности требуют либо выбора подходящего алгоритма определения особых точек, либо введения дополнительных средств, облегчающих анализ фото, например, добавления на кожу маркеров. Однако даже с дополнительными точками, которые обеспечат нам маркеры, модель, полученная таким образом, останется достаточно грубой. Для решения данной проблемы было принято решение получить дополнительные фото с помощью нахождения независимых промежуточных моделей отдельно по каждому фото. Таким образом, при решении поставленной задачи будут совмещены два кардинально различных подхода к получению пространственной модели по фото.

Алгоритм трехмерной реконструкции объекта с описанными выше особенностями можно разбить на следующие этапы:

1. Калибровка камеры.
2. Построение карты глубины для каждого изображения (частная трёхмерная модель из одного изображения).
3. Нахождение особых точек на каждом изображении.
4. Сопоставление особых точек.
5. Нахождение трёхмерных координат особых точек.
6. Объединение частных трёхмерных моделей по особым точкам.
7. Триангуляция с целью получения полигонального представления модели.

На рис. 1 представлен обобщённый алгоритм трёхмерной реконструкции, используемый в данной работе:

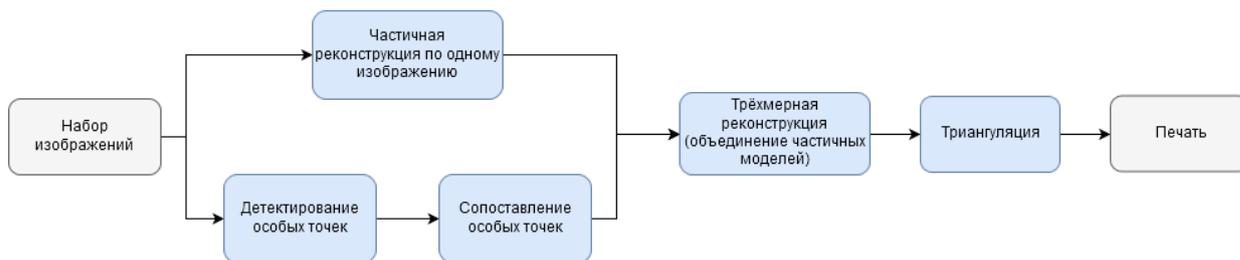


Рис. 1: Алгоритм трёхмерной реконструкции объекта обтекаемой формы

Обзор литературы

Задача трёхмерной реконструкции уже давно интересует специалистов. В зависимости от поставленных целей, специфики объектов реконструкции и имеющихся исходных данных, исходная задача разбивается на множество специфических.

Первым предложенным решением, которое, к слову, и по сей день является одним из самых широко используемых, является подход, основанный на выявлении особых точек (точек, выделяющихся в пределах своей окрестности и сохраняющихся при аффинных преобразованиях) и вычислений из трёхмерных координат на основе положений этих точек на фото или видео с разных ракурсов. Данный подход получил название Structure from Motion [1]. Так как в данном подходе ключевым моментом является оценка положения точки в пространстве на основе известных положений точки на нескольких изображениях, основополагающей работой для него является [2]. В ней собрана вся базовая информация по многомерной геометрии. Гибкость вышеописанного подхода осуществляется за счёт различных алгоритмов выбора особых точек. В [3] проведён сравнительный анализ некоторых детекторов особых точек. Если необходима высокая скорость детек-

тирования, то лучшим выбором будет ORB [4] и BRISK [5], SIFT [6] более устойчив к повороту, изменению насыщенности, зашумлению, перспективным искажениям и изменению масштаба, однако его скорость существенно ниже аналогов. SURF [7], являющийся упрощённой модификацией SIFT, работает гораздо быстрее своего прототипа, но менее устойчив к различного рода искажениям.

Вышеописанный подход, базирующийся на особых точках, прекрасно подходит для реконструкции больших сцен, архитектурных сооружений или просто предметов, содержащих большое число углов. Почему именно углов? Дело в том, что именно углы обеспечивают резкий перепад яркости на изображении, что в конечном счёте даёт особые точки. По этой причине детекторы особых точек также называют детекторами углов. У объектов обтекаемой формы с однородно окрашенной поверхностью, например, шар или ваза, внутренних особых точек, связанных с формой, обычно нет, а граничные точки не с чем сопоставлять. Проблему можно решить добавлением текстуры или разметки, но не для всех обтекаемых объектов это возможно. Очевидно, что применение одной лишь реконструкции по особым точкам в случае восстановления модели предплечья, которое является ярким примером объекта обтекаемой формы, не даст достаточно точных и адекватных результатов.

Однако не всегда есть возможность получить несколько изображений одного объекта. В этом случае рассматриваются подходы к восстановлению трёхмерной модели из единственного изображения. Очевидно, что информации, содержащейся на одном изображении недостаточно, чтобы получить трёхмерную модель. В этом случае необходима дополнительная информация об объекте реконструкции или о сцене. В зависимости от этой информации подходы к реконструкции делятся на несколько групп:

Перспективные. Методы основываются на информации о перспективе и том факте, что параллельные прямые сходятся в одной точке, а все точки схода лежат на одной линии – линии горизонта. Данный подход лучше всего подходит для восстановления больших сцен с большим числом параллельных линий (здания, городские улицы, коридоры, интерьер). В [9] рассматривается применение метода для реконструкции внутренних помещений. Для примера выбраны именно внутренние помещения (комната, коридор), так как при их изображении получают в основном одноточечную перспективу, что упрощает многие вычисления. Так как подход основывается на прямых параллельных линиях, очевидно, что он не применим для решения задачи реконструкции предплечья.

Силуэто-ориентированные. Целую серию статей проблеме трёхмерной реконструкции по единственному изображению посвятил Тёрре с соавторами. Их подходы относятся к силуэто-ориентированным, так как связывают силуэт с объёмом. В [10] рассматривается подход, согласно которому изображение сначала сегментируется, а затем по найденному силуэту находится минимальная выпуклая оболочка заданного объёма. Недостатком этого метода является зависимость корректности результатов от сегментации. В [11] данная идея продолжена, однако в данной работе вместо точечного представления модели используется воксельное. Происходит процесс наращивания модели до достижения заданного объёма. Как и в прошлом методе требуемый объём необходимо задавать вручную. В [12] при релаксации выпуклой оболочки авторы дают возможность выбора приоритета: приоритет формы – при релаксации стремятся к достижению оболочкой глобального минимума, – и приоритет объёма, когда допускается остановка итераций до достижения глобального минимума. В [13] авторы отказываются от метода выпуклой релаксации на каждой итерации, как в предыдущей работе, вме-

сто этого используется модификация градиентного спуска с применением пост-релаксации каждые n итераций.

Силуэто-ориентированные методы часто требуют предварительного задания объёма восстанавливаемой модели. Однако при применении данного подхода к реконструкции предплечья это не является проблемой, так как возможно предварительно собрать статистические данные, разбить их на группы, и для каждой группы вычислить средний объём предплечья. Таким образом, данный подход позволит максимально приблизить модель к реальному объекту.

Основанные на текстурах. В [14] описывается метод оценки формы объекта на основании оценки искажения равномерной текстуры. Для текстуры строится карта искажения текстуры, на основе которой затем вычисляется ориентация поверхности и форма (искривление поверхности). Как было сказано ранее, метод работает только при условии равномерной текстуры объекта, что делает его неприменимым для решения задачи, поставленной в данной работе.

Основанные на затенении. Метод берёт за основы понятие окклюзии – явления, когда объект за счёт своей формы сокращает доступ света в определённые области. Иными словами, это более широкое понятие собственной тени. Алгоритм, описанный в [15], вычисляет границу окклюзии и, основываясь на этой границе и на интенсивности пикселей изображения, позволяет оценить ориентацию и форму объекта. Методы, основанные на затенении, в отличие от методов, основанных на текстурах, предназначены для анализа однородно окрашенных объектов. Данный подход в достаточной мере пригоден для решения поставленной задачи трёхмерной реконструкции предплечья. Однако, так как изначально решение задачи было необходимо для печати лангеты для фиксации конечности при травме, изменения цвета

кожных покровов могут негативно сказаться на результатах реконструкции данным методом.

Использование скрытой марковской модели. Данный подход описывается в [16]. Процесс реконструкции делится на два этапа: обучение с учителем и непосредственно реконструкция. При обучении с учителем находится связь между интенсивностью пикселя и его глубиной в итоговой модели. Метод хорошо показывает себя при реконструкции обтекаемых объектов и чаще всего используется при реконструкции лица.

Помимо работ, описывающих базовые подходы, были рассмотрены прикладные работы, которые оценивали применение определённых методов к решению конкретных задач. В [17] рассмотрена мобильная реализация алгоритма трёхмерной реконструкции с помощью особых точек. Помимо стандартных техник в работе активно используются данные, получаемые с гироскопа и акселерометра, которыми сейчас оснащены большинство смартфонов. В [18] рассматривается применение технологии Structure from Motion в сфере судебной медицины для документирования процесса разложения. В данной работе также использовался подход, основанный на сопоставлении особых точек. В течение месяца снимки выполнялись через определённые промежутки времени, после чего все полученные модели были объединены для получения трёхмерной анимации процесса разложения. В [19] описано применение технологии Structure from Motion в геоморфологии, что позволило обойтись без дорогостоящего оборудования при получении структуры земных слоёв.

Глава 1. Реконструкция на основе одного изображения

Для реконструкции по единственному изображению был выбран метод, основывающийся на силуэте объекта. Достоинствами данного метода являются:

- Минимальные требования пользовательского ввода: задание одного параметра V_t .
- Скорость вычислений.
- Полное соответствие полученной модели силуэту объекта.

Недостатками же являются:

- Необходимость симметрии объекта относительно силуэта. В ходе реконструкции модель раздувается только в одну сторону относительно плоскости изображения.
- Невозможность пользователя оказывать влияние на форму восстанавливаемого объекта.
- Невозможность моделирования складок на поверхности объекта.

Для выбранной предметной области указанные недостатки метода никак не влияют на конечный результат, так как даже реконструкция производится автоматически, что исключает знание пользователя о промежуточных результатах реконструкции, а, следовательно, и возможность модификации им промежуточных моделей. Так как производится реконструкция предплечья, то также исключается наличие на восстанавливаемой поверхности складок. Условие симметричности вообще можно отбросить, так как на следующем

этапе работа проводится с частичной моделью (с "раздутой"половиной), а не с целой симметричной. Таким образом, метод как нельзя лучше подходит для применения в рамках данной предметной области.

Определение 2. Пусть $\Omega \in R^2$ - плоскость изображения. *Силуэтом объекта* будем называть множество точек $\Sigma \subset \Omega$, принадлежащих изображению объекта.

Определение 3. *Поверхностью минимального веса* будем называть множество точек $S \subset R^3$, где $\pi(S) = \Sigma$, $\pi : R^3 \rightarrow \Omega$, такое что

$$S = \operatorname{argmin}_S \int g(s) ds,$$

где $g(s) : V \rightarrow R_+$ - сглаживающая функция, где $V \subset R^3$ - область реконструкции.

Представим S в неявном виде, введя индикаторную функцию $u : V \rightarrow \{0, 1\}$, где $V \subset R^3$ - область реконструкции.

$$u(x) = \begin{cases} 0, & \pi(x) \notin \Sigma \\ 1, & x \in \Sigma \end{cases},$$

Индикаторные функции, соответствующие поверхностям, содержащим в себе силуэт Σ , образуют множество U_Σ . Зафиксируем желаемый объём V_t поверхности S . Индикаторные функции, соответствующие поверхностям S объёма V_t , образуют множество U_V :

$$U_V = \left\{ u \mid \int u(x) d^3x = V_t \right\}$$

Задача трёхмерной реконструкции по одному изображению принимает вид:

$$\min_{u \in U_\Sigma \cap U_V} E(u),$$

где

$$E(u) = \int g(x) |\nabla u(x)| d^3x,$$

При решении задачи необходимо выполнение трёх ограничений на u : согласованности с силуэтом, сохранении постоянного объёма V_t и $u \in [0, 1]$. Для обеспечения согласованности с силуэтом выполняется обновление только тех вокселей, которые находятся внутри контура силуэта. Для обеспечения выполнения двух других условий используется метод Дейкстры проекции на выпуклые множества. На каждой итерации вычисления производятся отдельно для объёма (1) и диапазона (2):

$$\begin{cases} u_V^i &= u_R^{i-1} - v_V^{i-1} + \frac{V_d}{N} \\ v_V^i &= u_V^i - (u_R^{i-1} - v_V^{i-1}) \end{cases}, \quad (1)$$

$$\begin{cases} u_R^i &= \Pi_{[0,1]}(u_V^i - v_R^{i-1}) \\ v_R^i &= u_R^i - (u_V^i - v_R^{i-1}) \end{cases}, \quad (2)$$

где $V_d = V_t - (u_R^{i-1} - v_V^{i-1})$, $\Pi_{[0,1]}(x)$ – проекция x на $[0, 1]$, N – число вокселей в заданном объёме V_t . Начальные значения: $u_R^0 = u^k$, $v_R^0 = 0$, $v_V^0 = 0$.

Использование метода градиентного спуска для нахождения глобального минимума может занимать много времени. Для сокращения времени вычислений используется метод выпуклой релаксации [20]:

$$\min_{u \in U_\Sigma \cap U_V} E(u) = \min_{u \in U^r} \sup_{\|\xi(x)\|_2 \leq g(x)} \left\{ \int -u \operatorname{div} \xi d^3x \right\},$$

где $\xi \in C^1(R^3, R^3)$, U^r – выпуклая оболочка $U_\Sigma \cap U_V$.

Оптимизация производится чередованием градиентного спуска по u и градиентного восхождения по ξ . Значения переменных на каждой итерации вычисляются следующим образом:

$$\begin{cases} \xi^{i+1} &= \Pi_{\|\xi(x)\|_2 \leq g(x)}(\xi^i + \tau \nabla \bar{u}^i) \\ u^{i+1} &= \Pi_{U^r}(u^i + \sigma \operatorname{div} \xi^{i+1}) \\ \bar{u}^{i+1} &= 2u^{i+1} - u^i \end{cases},$$

где $\Pi_{\|\xi(x)\|_2 \leq g(x)}(x)$ – проекция x на множество $\|\xi(x)\|_2 \leq g(x)$, U^r – выпуклая оболочка $U_\Sigma \cap U_V$.

На рис. 2 представлены результаты трёхмерной реконструкции отдельно для некоторых ракурсов:

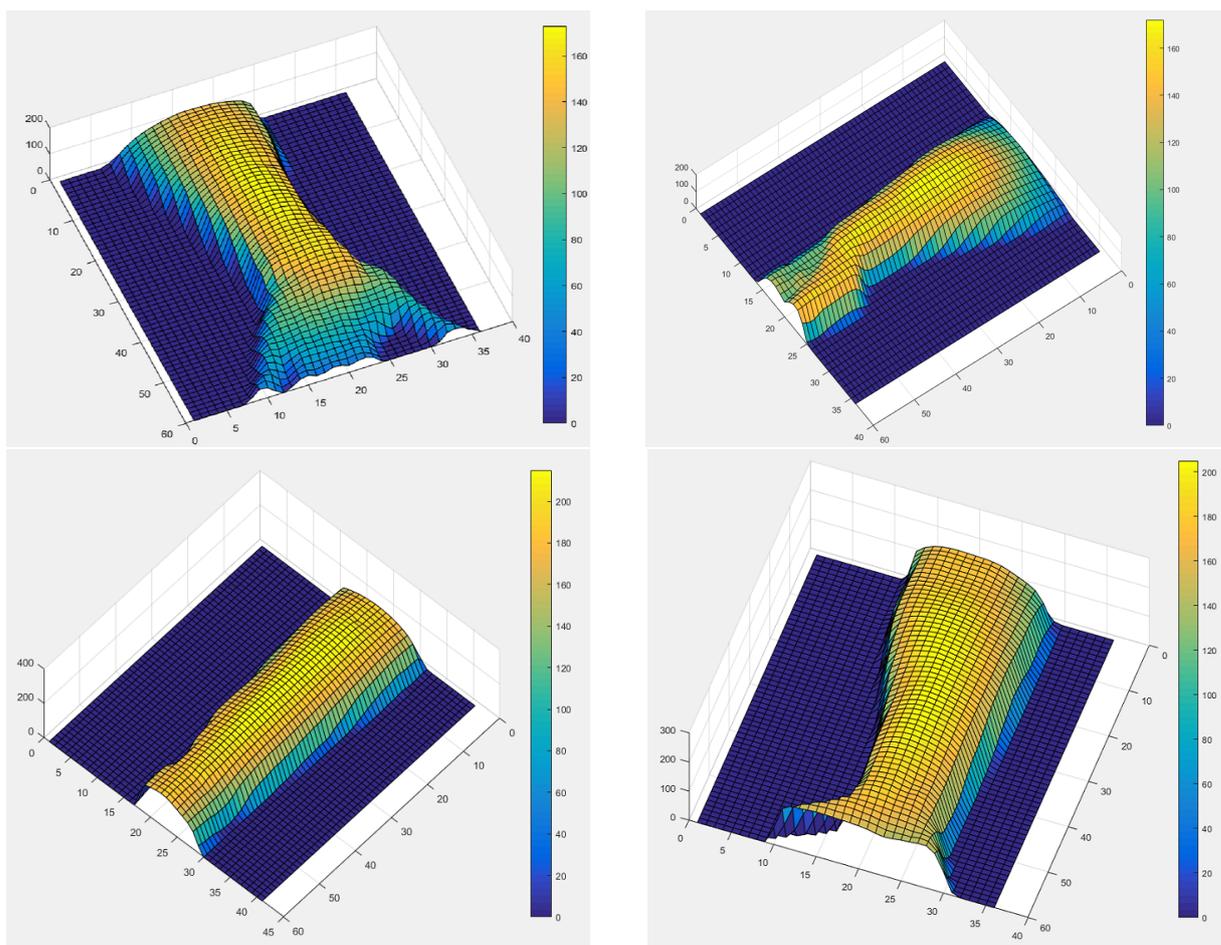


Рис. 2: Частичные трёхмерные модели построенные для разных ракурсов.

Глава 2. Особые точки (interested points)

Определение 4. *Особыми точками* называются точки на изображении, обладающие следующими свойствами:

- **Характерность** – точка должна отчётливо выделяться на фоне своей окрестности.
- **Инвариантность** – аффинные преобразования не должны оказывать влияние на точку.
- **Устойчивость** – точка должна быть устойчива к шумам на изображении.
- **Уникальность** – особая точка должна отличаться от других особых точек.
- **Интерпретируемость** – особые точки должны быть представлены таким образом, чтобы было возможным получение о них информации.

В [21] дополнительно вводятся следующие свойства особых точек:

- **Локальность** – при описании точки необходимо рассматривать очень небольшую её окрестность.
- **Численность** – число особых точек, обнаруженных на изображении, должно быть достаточным для работы с ними.
- **Точность** – соответствующие точки, обнаруженные на двух изображениях должны быть корректно сопоставимы.
- **Эффективность** – особые точки должны быть находимы в режиме реального времени или близком к нему.

2.1. Детекторы особых точек

Определение 5. *Детектором особых точек* называются точки или локальные области, имеющие некие признаки, выделяющие их среди множества соседних точек. Чаще всего особыми точками являются углы и резкие перепады освещенности.

Определение 6. *Дескриптором особой точки* называется функция, описывающая особую точку по её окрестности. Можно считать его альтернативным представлением особой точки, для которого, к тому же, можно ввести меру схожести.

2.1.1. ORB

ORB (Oriented FAST and Rotated BRIEF) – метод поиска и описания особых точек, в основе которого лежит комбинация детектора особых точек FAST и бинарного дескриптора BRIEF.

Детектор FAST

Классическая реализация детектора FAST (features from accelerated segment test) очень проста. FAST не производит лишних вычислений, например, значений градиентов яркости, а просто производит проверку яркости точек на изображении.

Для каждой точки-кандидата p рассматривается окружность с центром в этой точке. Обычно выбирается окружность из 16 точек. Значение яркости каждой точки на окружности сравнивается со значением яркости централь-

ной точки p :

$$S_x = \begin{cases} b, & I_x \geq I_p + t \\ s, & I_p - t \leq I_x \leq I_p + t, \\ d, & I_x \leq I_p - t \end{cases}$$

где I_p – значение яркости точки p , I_x – значение яркости точки x , принадлежащей окрестности p , t – пороговое значение яркости.

$$\begin{cases} P_b & = \{x \mid S_x = b\} \\ P_d & = \{x \mid S_x = d\} \end{cases}$$

Если $|P_b| = n$ или $P_d = n$, где обычно $n = 9$, то осуществляется проверка четырёх равноотстающих друг от друга точек на окружности. Если значение яркости ровно трёх из этих точек больше или меньше I_p , то точка считается особой.

Классическая реализация детектора FAST неинвариантна относительно поворота. В ORB используется модификация, обеспечивающая данную инвариантность. Для этого вводится параметр $\theta = \arctan(m_{01}, m_{10})$, отвечающий за ориентацию патча – области изображения, содержащей текущую точку-кандидата p с её окружностью $\rho(p)$. Где m_{01}, m_{10} – моменты яркости:

$$m_{mq} = \sum_{(x,y) \in \rho(p)} x^m y^q I(x, y).$$

Дескриптор BRIEF

Дескриптор BRIEF является бинарным дескриптором, то есть описывающим точку в виде бинарной строки длиной 256 бит (BRIEF-дескриптор в общем случае может быть описан 128, 256 и 512 битами, но в ORB стандартно требует использование 256 бит). Окрестность изображения разби-

вается на патчи – в данном случае блоки $N \times N$. Для каждого патча P проводится тест:

$$\tau(P, x, x') = \begin{cases} 1, & I(x) < I(x') \\ 0, & I(x) \geq I(x') \end{cases}$$

n тестов формируют матрицу $S_{2 \times n}$, которая содержит координаты протестированных пикселей. Для обеспечения инвариантности относительно поворота вычисляют ориентированную матрицу S_θ :

$$S_\theta = R_\theta S,$$

где R_θ – матрица поворота на угол θ .

Дескриптор особой точки p определяется как:

$$f_n(I) = \sum_{1 \leq i \leq n} 2^{i-1} \tau(P, x_i, x'_i)$$

ORB инвариантен относительно поворота, но не инвариантен относительно изменения масштаба, что не является существенной проблемой для решения задачи в том виде, в котором она была поставлена. Однако детектор FAST – это детектор углов, поэтому использование его для детектирования особых точек на гладких поверхностях может быть недостаточно успешно.

2.1.2. BRISK

BRISK (Binary Robust Invariant Scalable Keypoints) – бинарный дескриптор, устойчивый к повороту и изменению масштаба. Так же, как и у ORB, в основе него лежит использование комбинации детектора особых точек

FAST и бинарного дескриптора BRIEF.

Детектор

Для обеспечения инвариантности к масштабированию вводится пространство масштаба. Для этого строится пирамида, состоящая из n октав c_i и $n - 1$ внутренних октав d_i . Обычно берётся $n = 4$. При этом октаву c_i получают путём сжатия исходного изображения c_0 в 2^i раз, а d_i – в $1.5 \cdot 2^i$ раз. На каждой октаве особые точки вычисляются FAST-детектором. Точки-кандидаты исходного изображения интерполируются на каждую октаву и среди этих соответствующих точек выбирается точка с максимальным значением интенсивности.

Дескриптор

Дескриптор строится следующим образом. Область вокруг особой точки разбивается на N участков. На рис. 3 приведена схема для $N = 60$.

Пусть (p_i, p_j) – пара рассматриваемых точек в окрестности A особой точки p , где $A = \{(p_i, p_j) \in R^2 \times R^2 | i < N; j < i; i, j \in N\}$.

Множество A разбивается на два подмножества:

$$L = \{(p_i, p_j) \in A | \|p_j - p_i\| > \rho_{min}\}$$

$$S = \{(p_i, p_j) \in A | \|p_j - p_i\| < \rho_{max}\}$$

где ρ_{min} , ρ_{max} – пороговые значения расстояния между точками, равные соответственно $\rho_{min} = 13.67t$, $\rho_{max} = 9.75t$, где t – масштаб особой точки.

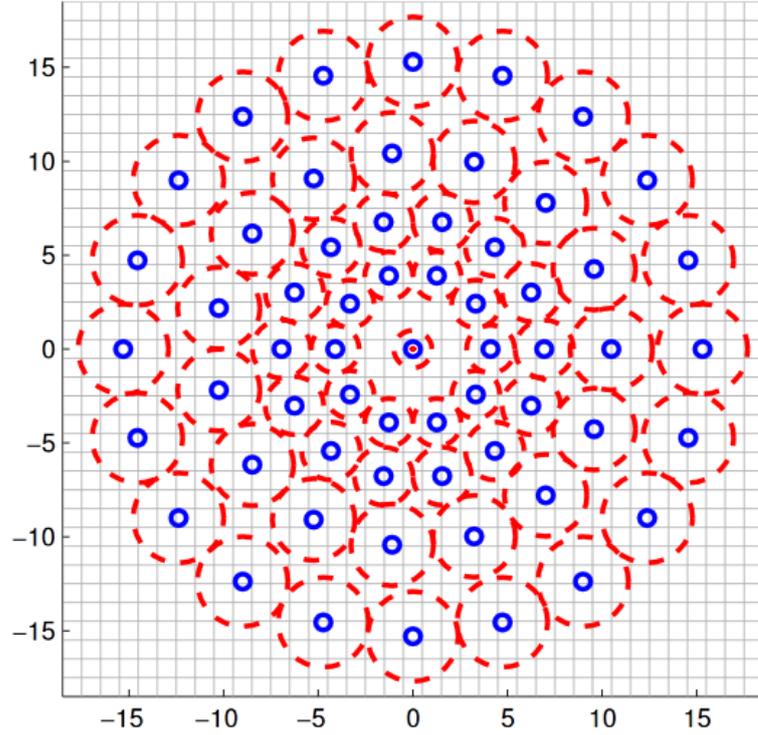


Рис. 3: Шаблон окрестности особой точки p для вычисления дескриптора при $N = 60$ ¹.

Вычисляется среднее значение градиента:

$$g = \begin{pmatrix} g_x \\ g_y \end{pmatrix} = \frac{1}{|L|} \sum_{(p_i, p_j) \in L} (p_j - p_i) \frac{I(p_j, \sigma_j) - I(p_i, \sigma_i)}{\|p_j - p_i\|},$$

где $I(p_i, \sigma_i)$ – интенсивность точки p_i после размытия по Гауссу с ядром σ_i . Размытие по Гауссу с размером ядра σ пропорциональным расстоянию от данной точки окрестности до центральной особой точки p используется, чтобы устранить эффект наложения при отображении яркости точек на шаблон для них.

Среднее значение градиента используется для определения направления $\theta = \arctan(g_y, g_x)$.

Далее строится BRIEF-дескриптор на множестве $S_\theta = R_\theta S$.

¹Изображение взято из BRISK: Binary Robust Invariant Scalable Keypoints [5].

2.1.3. SIFT

Детекторы SIFT и SURF, которые будут описаны далее, относятся к детекторам блобов.

Определение 7. *Блоб* – это область изображения, в которой некоторые свойства являются постоянными или приблизительно постоянными. Все точки в блобе можно в каком-то смысле считать похожими друг на друга.

Алгоритм SIFT (Scale-Invariant Feature Transform) устойчив к повороту, изменению масштаба, изменению яркости и перспективным преобразованиям.

Детектор

Алгоритм поиска особых точек можно разбить на несколько этапов:

1. Представление изображения в пространстве масштаба.

Пусть $I(x, y)$ – входное изображение, тогда

$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$ – его представление в пространстве масштаба, где $G(x, y, \sigma)$ – фильтр Гаусса:

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

Строится пирамида, состоящая из n октав с коэффициентами размытия прямопропорциональными масштабу.

2. Построение разности гауссианов. DoG называют разность между двумя последовательными гауссианами:

$$DoG = G(x, y, k\sigma) - G(x, y, \sigma) = (k - 1)\sigma^2 \nabla^2 G \quad (3)$$

Максимумы и минимумы DoG (3) соответствуют устойчивым признакам в пространстве масштаба.

На рис. 4 представлена пирамида гауссианов и получаемая из неё пирамида DoG.

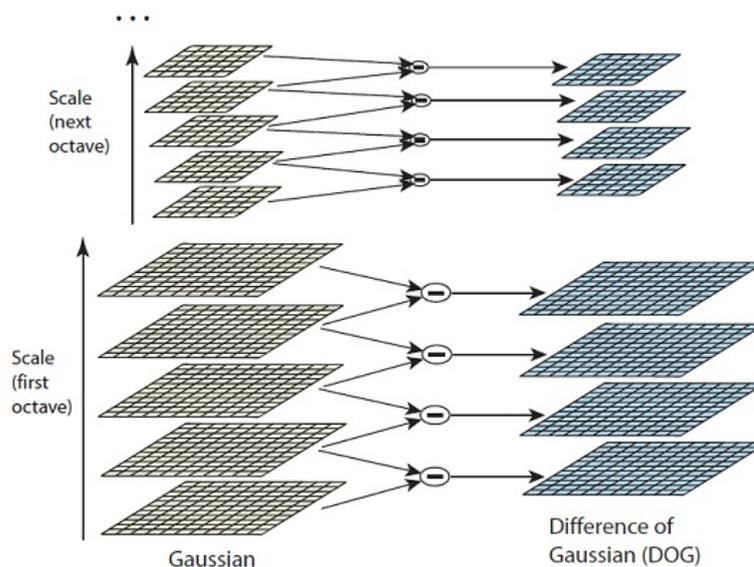


Рис. 4: Пирамида гауссианов и соответствующая ей пирамида их разностей.²

3. На каждой октаве DoG ищутся локальные максимумы и минимумы – точки-кандидаты.
4. Из найденных на шаге 3 точек-кандидатов отсекаются точки, в которых экстремум слабо выражен и граничные точки.

Дескриптор

SIFT-дескриптор представляет собой гистограмму градиентов в окрестности особой точки.

Построение гистограммы градиентов производится следующим образом. В каждой точке окрестности находится величина градиента ∇I и его

²Изображение взято из SIFT: Theory and Practice. <http://aishack.in/tutorials/sift-scale-invariant-feature-transform-log-approximation/>.

направление θ . Строится гистограмма ориентаций. Для этого пространство возможных направлений разбивается на 36 корзин по 10° . Пусть H_i – i -я компонента гистограммы, соответствующая направлению $[i \cdot 10^\circ; (i+1) \cdot 10^\circ)$, тогда величина этой компоненты равна:

$$H_i = \sum_{\theta \in [i \cdot 10^\circ; (i+1) \cdot 10^\circ)} |\nabla I| G(x, y, \sigma).$$

В итоге особой точке присваивается направление θ , соответствующее максимуму в построенной гистограмме.

2.1.4. SURF

Метод SURF (Speeded Up Robust Features) является модификацией метода SIFT, обладающей более высокой скоростью работы. Так же, как и он, SURF инвариантен относительно поворота, изменения масштаба, изменения яркости и перспективным преобразованиям.

Детектор

Для обеспечения инвариантности относительно масштаба, как и в SIFT используется пирамида октав. Для аппроксимации лапласиана гауссианов вместо разности гауссианов используется матрица Гессе, детерминант которой достигает экстремума в точках максимального изменения яркости. Матрица Гессе имеет вид:

$$H(I(x, y)) = \begin{pmatrix} L_{xx}(x, y, \sigma) & L_{xy}(x, y, \sigma) \\ L_{xy}(x, y, \sigma) & L_{yy}(x, y, \sigma) \end{pmatrix},$$

где $L_{xx} = I(x, y) * \frac{\partial^2 G(x, y, \sigma)}{\partial x^2}$, $L_{xy} = I(x, y) * \frac{\partial^2 G(x, y, \sigma)}{\partial x \partial y}$.

Её детерминант называется гессианом и имеет вид:

$$\det H = D_{xx}D_{yy} - (\gamma D_{xy})^2,$$

где D_{xx} , D_{yy} , D_{xy} – аппроксимации лапласиана L_{xx} , L_{yy} , L_{xy} ,

$\gamma = 0.9$ – коэффициент, компенсирующий приближенность вычислений.

На рис. 5 представлены дискретные фильтры для нахождения L_{xx} , L_{yy} , L_{xy} , а на рис. 6 их аппроксимация для нахождения D_{xx} , D_{yy} , D_{xy} .

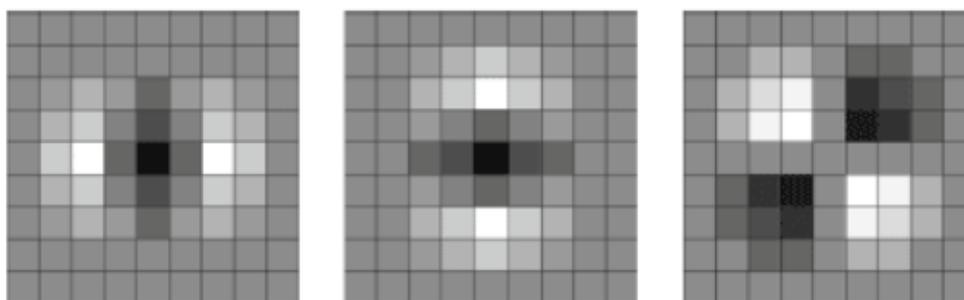


Рис. 5: Дискретные фильтры для нахождения L_{xx} , L_{yy} , L_{xy} .

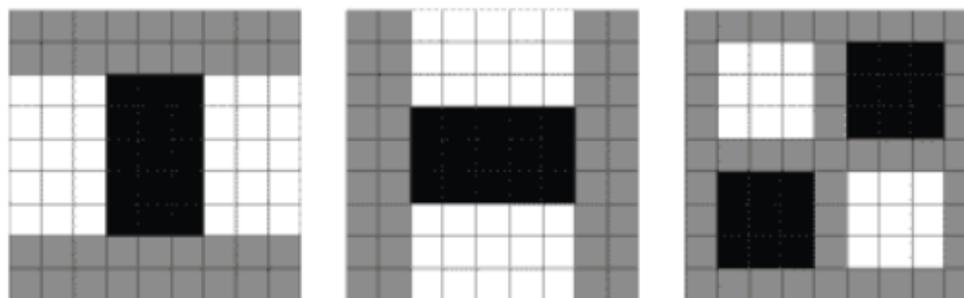


Рис. 6: Аппроксимация дискретных фильтров для нахождения D_{xx} , D_{yy} , D_{xy} .

Для определения ориентации особой точки используется фильтр Хаара. Вычисляются взвешенные значения изменения яркости по x и по y , на основе которых определяется ориентация особой точки.

Дескриптор

Для построения дескриптора также используется фильтр Хаара размером 2σ , где σ – масштаб найденной особой точки. Область вокруг осо-

бой точки разбивается на 16 блоков, которые, в свою очередь, разбиваются ещё на 16 подблоков. Для каждого подблока находится вектор градиента (dx, dy) , который поворачивается на угол ориентации особой точки. В итоге особая точка описывается четвёркой значений $(\sum dx, \sum |dx|, \sum dy, \sum |dy|)$.

2.2. Сопоставление особых точек

2.2.1. Метрики

Пусть $descr_i$ – дескриптор особой точки p_i первого изображения, $descr'_j$ – дескриптор особой точки p'_j второго изображения. Для сопоставления особых точек необходимо введение метрики $\rho(descr_i, descr'_j)$, дающей численную характеристику меры похожести одного дескриптора на другой. В зависимости от типа дескриптора выбираются разные метрики:

1. Для бинарных дескрипторов, к которым относятся дескрипторы BRIEF, BRISK и ORB в качестве метрики обычно используется расстояние Хэмминга. Данное расстояние вычисляется как число несовпадающих символов на соответствующих местах двух дескрипторов:

$$\rho(descr_i, descr'_j) = \sum_{k=0}^K I(descr_i(k) \neq descr'_j(k)),$$

где K – размерность дескриптора, $descr_i(k)$ – k -ый элемент вектора признаков $descr_i$.

Замечание 1. Расстояние Хэмминга в качестве метрики применимо не только к бинарным дескрипторам, но и к любым дескрипторам, представимым в виде строк k -символьного алфавита.

2. Для численных дескрипторов, к которым относятся SIFT и SURF, обыч-

но выбирают евклидово расстояние:

$$\rho(descr_i, descr'_j) = \sum_{k=0}^K \sqrt{(descr_i(k) - descr'_j(k))^2},$$

где K – размерность дескриптора, $descr_i(k)$ – k -ый элемент вектора признаков $descr_i$,

или манхэттенское расстояние:

$$\rho(descr_i, descr'_j) = \sum_{k=0}^K |descr_i(k) - descr'_j(k)|.$$

2.2.2. Алгоритмы сопоставления

1. Brute-Force Matcher

Данный алгоритм сопоставления для каждой точки p_i первого изображения перебирает все точки p'_j второго изображения, вычисляет расстояния между их дескрипторами $\|descr_i - descr'_j\|$ и на выходе выдаёт k "ближайших соседей" точки p_i на втором изображении. Всего в ходе сопоставления выполняется $N * M$ проверок, где N – число особых точек на первом изображении, M – число особых точек на втором изображении.

2. FLANN Matcher

FLANN Matcher (Fast Approximate Nearest Neighbor Search Library)[22] – библиотека быстрого приблизительного поиска ближайших соседей, доступная для C, C++, MATLAB и Python. Данный подход разработан специально для большого набора данных. Быстродействие достигается за счёт индексации дескрипторов и дальнейшей проверки не всех претендентов, а только лишь небольшого набора, составленного с учётом

индекса.

Так как одной из проблем трёхмерной реконструкции объекта обтекаемой формы было недостаточное число особых точек, очевидно, что для данной предметной области эффективнее будет использование Brute-Force Matcher.

2.2.3. Фильтрация ложных соответствий

Алгоритм фильтрации ложных соответствий имеет следующий вид:

1. Пусть $P_N = \{p_i\}$, $i = \overline{1, N}$, $P'_M = \{p'_j\}$, $j = \overline{1, M}$ – множества особых точек на первом и втором изображениях соответственно.

2. $descr_i$ – дескриптор особой точки $p_i \in P_N$

$descr'_j$ – дескриптор особой точки $p'_j \in P'_M$

3. Методом k ближайших соседей ставим в соответствие каждой точке p_i на первом изображении точки p'_{j^*} и $p'_{j^{**}}$ на втором, где

$$j_i^* = \operatorname{argmin}_{j=\overline{1, M}} \|descr_i - descr'_j\|,$$

$$j_i^{**} = \operatorname{argmin}_{j \in \{1, \dots, M\} \setminus j_i^*} \|descr_i - descr'_j\|.$$

4. $d_{j^*} = \|descr_i - descr'_{j^*}\|$, $d_{j^{**}} = \|descr_i - descr'_{j^{**}}\|$ – расстояния между дескриптором точки p_i первого изображения и дескрипторами точек p'_{j^*} и $p'_{j^{**}}$ второго изображения соответственно. $d_{j^*} < d_{j^{**}}$.

5. Если $d_{j^*} < kd_{j^{**}}$, где $k \approx 0.6 - 0.8$, то в список соответствий добавляем пару $p_i \leftrightarrow p'_{j^*}$.

Если $d_{j^*} \in [kd_{j^{**}}, d_{j^{**}}]$, то точку p_i отбрасываем.

Параметр k определяет необходимую для однозначного сопоставления степень отличия точки-претендента от точки, следующей за ней по похо-

жести. Для разных изображений и разных предметных областей k может варьироваться.

2.3. Реализация

На рис. 7 представлены результаты применения детекторов ORB, BRISK, SIFT и SURF к изображению, соответствующему поставленной задаче. Можно заметить, что большинство особых точек сконцентрировано на границе объекта (руки) и, что ещё более плохо, на фоне. Это демонстрирует необходимость нанесения дополнительных маркеров для возможности использования найденных особых точек при реконструкции.



a) ORB детектор



b) BRISK детектор



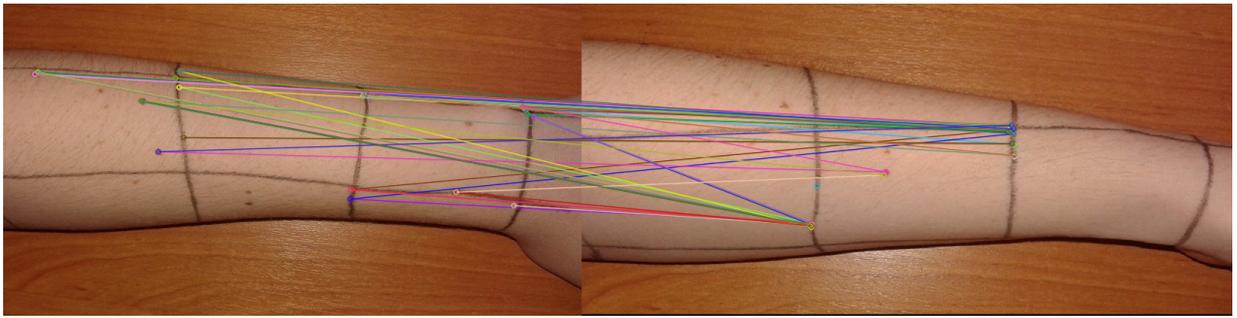
c) SIFT детектор



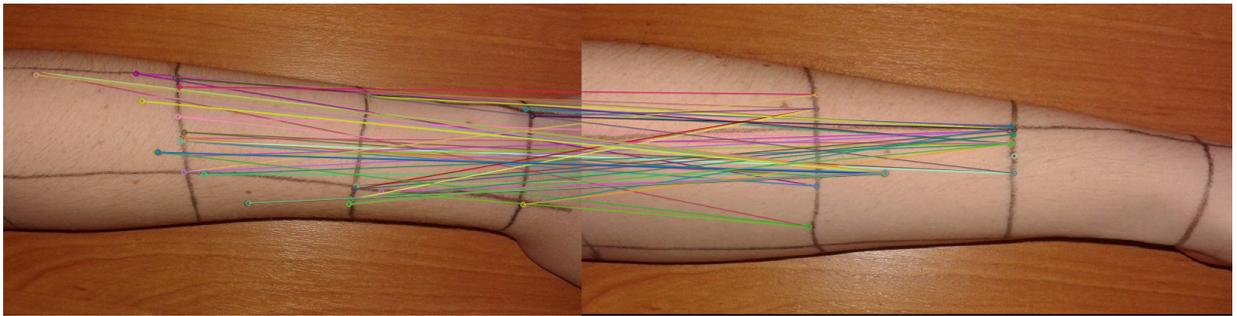
d) SURF детектор

Рис. 7: Результаты применения различных видов детекторов особых точек к изображению без маркеров.

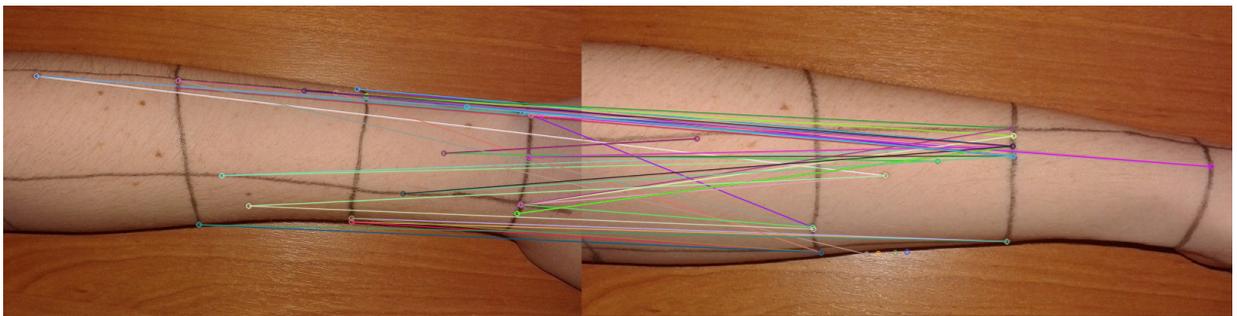
На рис. 8 приведены результаты детектирования и сопоставления особых точек на изображениях с нанесёнными маркерами.



a) ORB детектор



b) BRISK детектор



c) SIFT детектор



d) SURF детектор

Рис. 8: Результаты применения различных видов детекторов особых точек к изображению с маркерами и их последующего сопоставления.

Проанализировав результаты сопоставления, можно сделать вывод, что лучше всего при решении поставленной задачи показал себя метод SIFT, так как именно он дал меньше всего ложных соответствий.

Глава 3. Калибровка камеры

На рис. 9 представлена используемая в компьютерном зрении модель камеры.

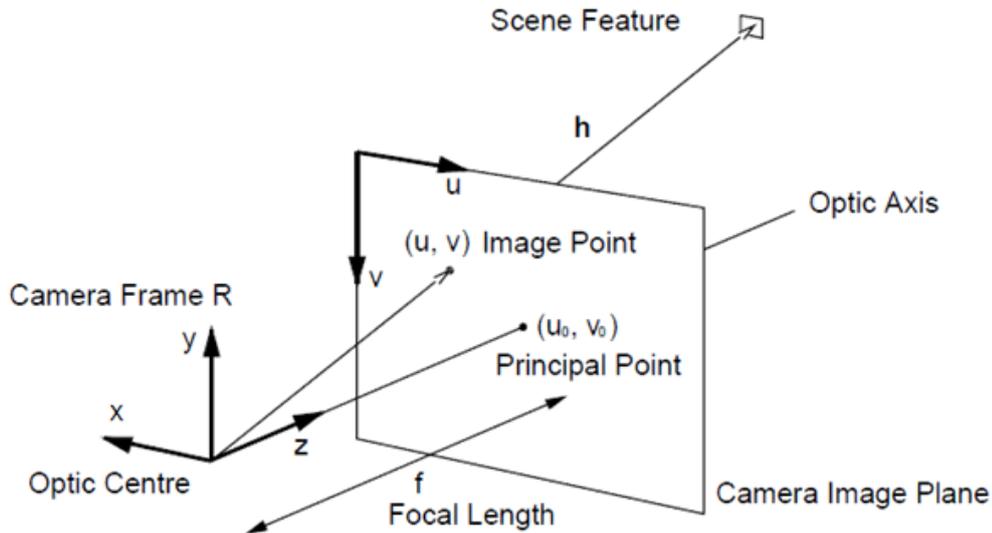


Рис. 9: Классическая модель камеры.³

Определение 8. *Внутренними параметрами камеры* называются параметры, связанные непосредственно с её техническими характеристиками. Внутренние параметры камеры задают калибровочную матрицу, имеющую следующий вид:

$$K = \begin{bmatrix} \alpha & -\alpha \operatorname{ctg} \theta & u_0 \\ 0 & \beta / \sin \theta & v_0 \\ 0 & 0 & 1 \end{bmatrix},$$

где $\alpha = k_x f$ и $\beta = k_y f$ – фокусное расстояние, выраженное в пикселях (k_x и k_y – разрешение матрицы камеры, $\frac{1}{k_x}$, $\frac{1}{k_y}$ – соответственно ширина и высота пикселя. Полагается, что допустимы не квадратные пиксели), u_0 и v_0 – координаты пересечения оптической оси с плоскостью изображения, θ – угол наклона (искажения) пикселя (у неискаженного пикселя $\theta = 90^\circ$).

Матрица внутренних параметров камеры K необходима для связи коор-

³Изображение взято из Multiple View Geometry in Computer Vision [2].

динат в плоскости изображения с координатами в пикселях.

Иногда производитель камеры указывает её внутренние характеристики, но если они неизвестны, их всегда можно получить посредством калибровки. Так как матрица калибровки зависит только от самой камеры, то достаточно однократно её найти, и в дальнейшем использовать во всех случаях, когда она необходима.

Определение 9. *Внешними параметрами камеры* называются параметры, характеризующие положение мировой системы координат относительно камеры. Иными словами, внутренними параметрами камеры является матрица поворота R и вектор переноса t .

Определение 10. *Калибровкой камеры* называется процесс получения внешних и внутренних характеристик камеры.

Для калибровки используют набор тестовых изображений с заданными характеристиками. Классическим примером калибровочных изображений является изображение шахматной доски с заранее известным числом клеток.

Замечание 2. Во всех случаях, когда говорится о координатах, имеется в виду, что речь идёт об однородных координатах, если явно не указано обратное.

Определение 11. *Однородными координатами* точки $M(a, b, c)$ в трёхмерном пространстве называется четвёрка чисел x, y, z, w , таких, что $a = \frac{x}{w}$, $b = \frac{y}{w}$, $c = \frac{z}{w}$, $x^2 + y^2 + z^2 + w^2 \neq 0$. Однородные координаты определяются с точностью до умножения на ненулевой множитель.

На практике калибровка камеры разбивается на два этапа: предварительная калибровка – получение внутренних параметров камеры, если они неизвестны, и определение положения камеры в пространстве (её внешних параметров). Матрица внешних параметров вычисляется непосредственно

в процессе трёхмерной реконструкции.

Глава 4. Объединение моделей

В отличие от получения кадров из видеопотока, в данном случае нельзя гарантировать "последовательность" изображений. В общем случае изображения могут поступать на обработку в произвольном порядке. По этой причине нельзя сразу выделить сообщающиеся кадры.

Замечание 3. *Сообщающимися кадрами* будем считать кадры, сделанные одной камерой последовательно друг за другом за достаточно малый отрезок времени Δt при движении по заданной траектории. Выполнение поставленных условий в большинстве случаев обеспечивает наличие на сообщающихся кадрах достаточного для реконструкции числа сопоставленных особых точек.

При отсутствии сообщающихся кадров выполняется сопоставление всех возможных пар изображений. Так как реконструкция осуществляется по конечному набору изображений, размер которого при постановке задачи ограничился четырьмя кадрами с различных ракурсов, необходимо будет перебрать всего 6 пар. В случае изменения числа входных кадров до n число пар становится C_n^2 .

Пусть на данном этапе для каждого входного изображения I_k , $k = \overline{1, n}$, $n = 4$, обнаружены соответствующие множества особых точек P^k . Алгоритм объединения моделей можно разбить на следующие шаги:

1. Обнаружение соответствий между особыми точками для всех возможных пар изображений.

$$\overline{P_{nm}} = \{(p_i^n, p_j^m) | p_i^n \in P^n, p_j^m \in P^m, n \neq m\}, n = \overline{1, N}, m = \overline{1, M}.$$

2. Для каждого изображения выбрать два других с наибольшим числом найденных соответствий.
3. Для отобранных пар изображений вычислить трёхмерные координаты особых точек по обнаруженным соответствиям.

Рассмотрим подробнее третий шаг. Его, в свою очередь, можно разбить на три основных шага:

1. Нахождение фундаментальной матрицы F .

Пусть даны I и I' – пара последовательных изображений, $x_i \leftrightarrow x'_i$ – пары соответствующих точек на них.

Пара точек x_i и x'_i является проекцией некоторой точки X_i мирового пространства тогда и только тогда, когда существует матрица $F_{3 \times 3}$, такая, что

$$x_i'^T F x_i = 0. \quad (4)$$

Для существования единственного решения уравнения (4) относительно F с точностью до умножения на константу, необходимо, чтобы в наличии имелось не менее 7 пар особых точек.

2. Нахождение матриц камер P и P' (определение их внешних параметров).

Так как матрицы камер имеют вид $P = K[R|t]$ и $P' = K'[R'|t']$, и предполагается, что матрицы внутренних параметров соответствующих камер K' , K известны, то целью данного шага является нахождение внешних параметров камер.

Возьмём $P = K[I|0]$, тогда t и R задают положение второй камеры

относительно первой. Вычислим существенную матрицу E :

$$E = K'^T F K = [t]_x R,$$

Вторая камера может иметь только четыре различных положения относительно первой:

$$\begin{aligned} P' &= K'[R_1|\gamma t] & P' &= K'[R_1|-\gamma t] \\ P' &= K'[R_2|\gamma t] & P' &= K'[R_2|-\gamma t] \end{aligned}$$

где t – решение уравнения $E^T t = 0$, R_1 и R_2 получены из E путём SVD-разложения, $\gamma > 0$ – масштабный коэффициент. Возможные относительные расположения камер представлены на рис. 10.

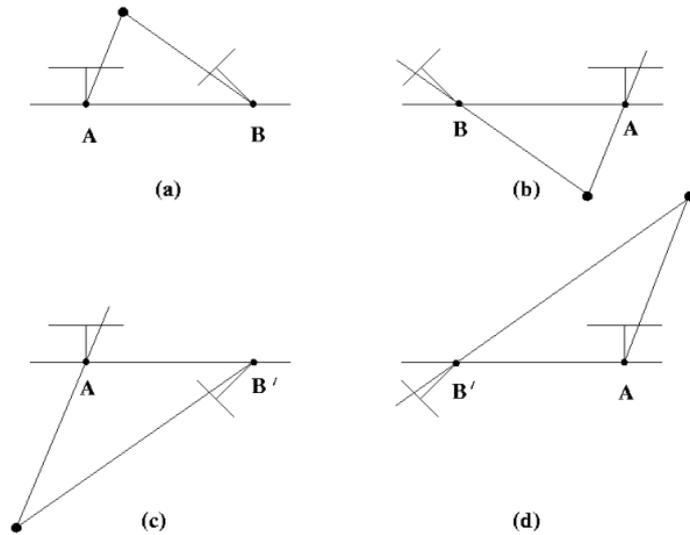


Рис. 10: Возможные относительные положения камер.

Однако только в одном случае из четырёх все найденные особые точки лежат в области видимости обеих камер.

3. Для каждой пары $x_i \leftrightarrow x'_i$ нахождение их проекции X_i в мировом пространстве.

Пусть X_i – искомая трёхмерная точка, проецируемая в точку $x_i \in I$ и в точку $x'_i \in I'$, P и P' – матрицы внешних параметров соответствующих камер. Тогда можно составить уравнение:

$$\begin{cases} p_i = PX_i \\ p'_i = P'X_i \end{cases}$$

Координаты искомой точки X_i получаем путём решения уравнения (3).

Имея пространственные координаты особых точек каждой частичной трёхмерной модели, можно получить аффинное преобразование для каждой точки частичной трёхмерной модели, в результате которого на выходе получаем точечное представление искомой модели.

На рис. 11 представлена точечная итоговая модель – результат объединения частичных моделей. Для наглядности на графике выведены не все точки модели, а лишь небольшое число равноотстающих узлов.

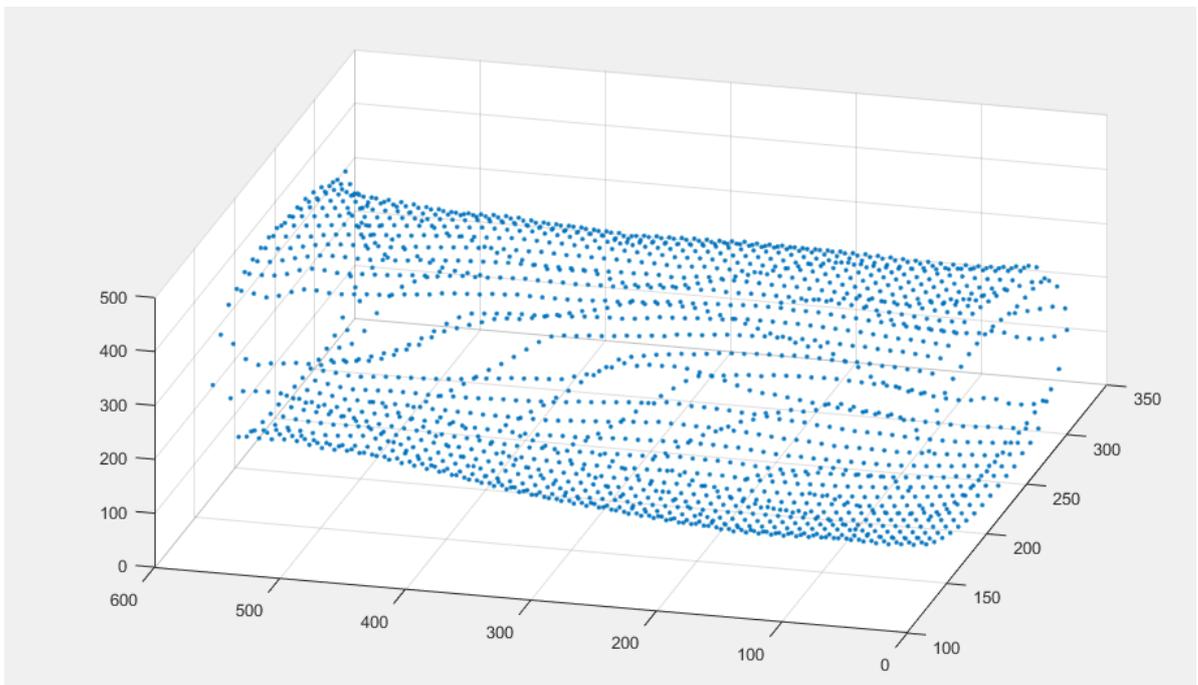


Рис. 11: Точечная модель, полученная в результате объединения частичных моделей.

Глава 5. Представление трехмерных моделей

Существует несколько способов представления моделей. Все их можно разделить на две группы: точные (непрерывные) и приближённые (дискретные).

- Дискретное представление

- Точечная модель

Данная модель представляет собой поверхность в качестве простого набора точек в трехмерном пространстве, не содержащем никакой информации кроме координат. Данное представление удобно для предварительного просмотра модели и является базовым для перехода к другим представлениям.

- Объемная (воксельная) модель

Воксель - аналог пикселя для трехмерного изображения. Иными словами, модель представляется как множество кубиков. Воксельная модель может быть как регулярной (все воксели одного размера), так и нерегулярной. Данное представление способно представить сложные формы простым образом, однако оно требует больших затрат памяти и плохо масштабируемо.

- Модель, основанная на изображениях

При данном представлении каждому пикселю изображения ставится в соответствие глубина - расстояние от данной точки на реальном объекте до камеры.

- Непрерывное (поверхностное) представление

При непрерывном представлении модель описывается как бесконечное множество точек, о каждой из которых информация известна. Как пра-

вило, непрерывные модели невозможно получить непосредственно при помощи сканирующего оборудования. Данное представление получают из дискретных моделей путем их анализа и различных преобразований.

– Сплайновая (аналитическая) модель

Данная модель представляет поверхность в виде особого набора функций - сплайнов, которые по опорным точкам интерполируют поверхность или отдельную ее часть.

– Векторно-полигональная модель

Данная модель, как и точечная, хранит набор вершин, однако помимо этого имеется дополнительная информация о связях между ними (гранях и полигонах). Кроме того, в зависимости от поставленных целей, можно хранить модель в виде:

1. набора граней, представленных из вершинами,
2. набора вершин и списка рёбер, содержащего указатели на входящие в них вершины,
3. набора вершин, списка рёбер с указателями на вершины и списка граней с указателями на входящие рёбра.

Для трёхмерной печати обычно используют формат STL (от "stereolithography"). Это формат, специально разработанный для технологий, производящих послойное наращивание объектов. Формат позволяет хранить модель в виде набора треугольных полигонов и их нормалей. Данное представление возможно получить из точечного или произвольного полигонального путём триангуляции. Формат впервые был представлен в 1987 году. Файл данного формата хранит для каждой вершины полигона и каждой нормали по три числа, таким образом, один полигон описывается двенадцатью числами. Стоит отметить, что данный формат с увеличением слож-

ности модели – а именно числа полигонов – быстро увеличивает объём файла.

После конвертирования модели в формат STL её передают на обработку слайсеру. Слайсер – это программа, которая принимает на вход STL-модель, "нарезает" её на слои, а на выходе выдает набор машинных инструкций для печати этих слоёв[23].

Глава 6. Программная реализация

Реализацию предложенного метода трёхмерной реконструкции можно условно разбить на два этапа: этап проведения вычислений и этап визуализации результатов. Для каждого этапа было разработано своё приложение. Связь между приложениями осуществлялась путём записи результатов вычислений в файл с расширением .dat первой программой и чтения из него второй.

- Этап вычислений

Приложение на данном этапе было написано на языке C++14 с использованием библиотеки OpenCV 3.3.1 с дополнительными модулями из opencv_contrib. Из дополнительных моделей были использованы модули sfm (Structure From Motion) и xFeatures2d.

- Этап визуализации

Приложение для визуализации было написано на языке MATLAB. Написание стороннего приложения для визуализации было обусловлено невозможностью отображения трёхмерных объектов средствами C++ и OpenCV и сложностью, наряду с отсутствием гибкости, при подключении других библиотек C++. MATLAB, напротив, предоставил весь

необходимый инструментарий для быстрого отображения результатов вычислений с возможностью манипулирования полученным представлением (например, возможностью приближения отображаемой модели и её поворота).

На изображениях 12, 13, 14, 15 приведены результаты работы программ на каждом из этапов реконструкции.

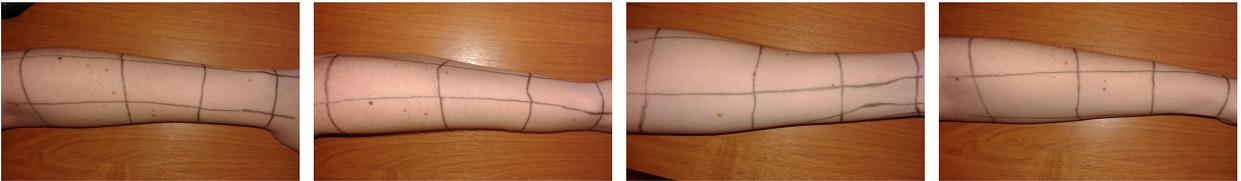


Рис. 12: Изображения, поданные на вход.

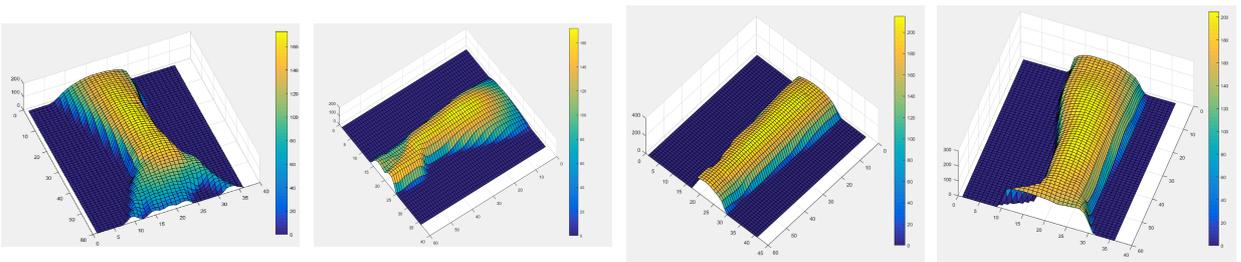


Рис. 13: Частичные трёхмерные модели, полученные каждая из одного изображения.

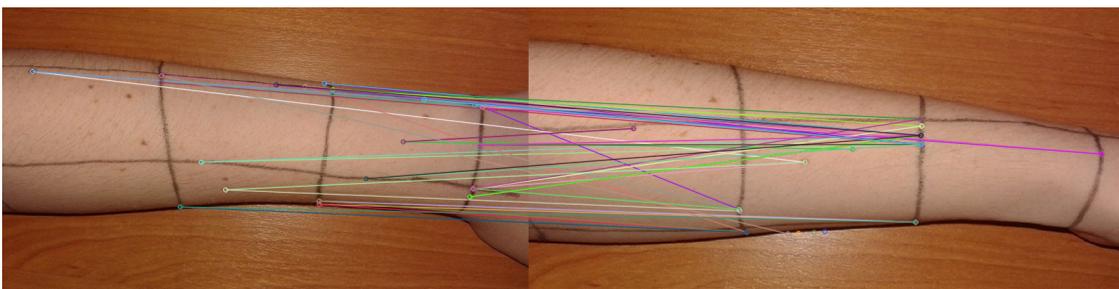


Рис. 14: Обнаружение и сопоставление особых точек на исходных изображениях (на примере одного фото).

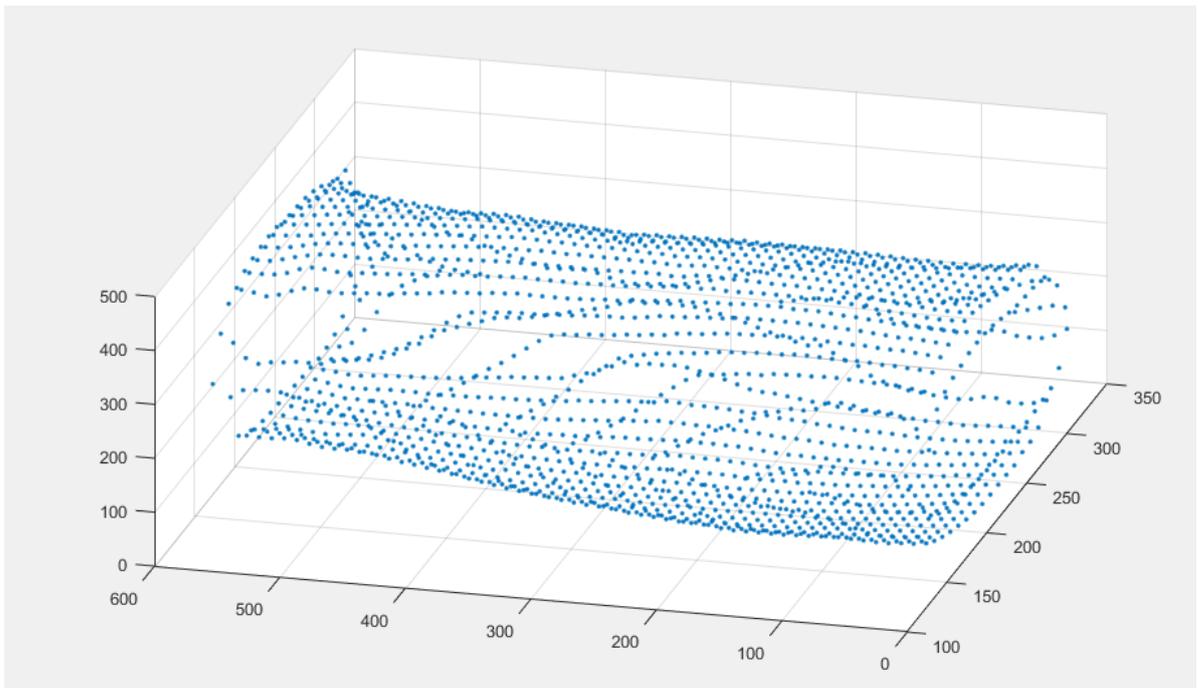


Рис. 15: Точечная финальная модель.

Выводы

В ходе решения поставленной задачи был выявлен ряд её особенностей:

1. Точная реконструкция гладких поверхностей на основании только лишь особых точек невозможна.
2. Использование на промежуточном этапе метода восстановления по одному изображению позволило значительно улучшить качество реконструкции.
3. На гладких поверхностях однородного цвета возникают проблемы с детектированием особых точек. Данная проблема решается нанесением дополнительной разметки.
4. При решении задачи нахождения и сопоставления особых точек в данной предметной области лучше всего себя показал метод SIFT, который

при достаточно большом числе найденных точек дал меньше всего ложных соответствий.

5. Градиентный спуск для решения задачи оптимизации поверхности заданного объёма при реконструкции из одного изображения сходится очень медленно. Существенный прирост скорости обеспечивает применение вместо него метода выпуклой релаксации.
6. После получения точечной модели для её печати необходимо провести триангуляцию для получения векторно-полигонального представления. Качество и скорость получения векторно-полигональной модели из точечной зависит от выбранного алгоритма триангуляции.

Полученная в итоге модель визуально соответствует объекту реконструкции, содержит в себе достаточное число полигонов, чтобы обеспечить требуемую точность представления гладкой поверхности. Модель пригодна для дальнейшей трёхмерной печати. Однако возможность практического применения подхода на данный момент невозможно оценить по причине отсутствия проведения тестов модели на людях.

Заключение

В данной работе был рассмотрен подход к трёхмерной реконструкции объектов обтекаемой формы. Реконструкция проводилась в два этапа:

1. Частичная трёхмерная реконструкция по одному изображению, которая выполнялась для каждого, подающегося на вход изображения.
2. Объединение частных моделей в одну общую.

На первом этапе для частичной реконструкции был рассмотрен подход, основанный на нахождении поверхности заданного объёма для заданного силуэта. Данный метод хорошо себя показывает при реконструкции гладких (здесь имеется в виду геометрический смысл понятия "гладкий") поверхностей.

На втором этапе были рассмотрены различные детекторы и дескрипторы особых точек. В ходе экспериментов было установлено, что для определения глобальных координат особых точек необходимо введение дополнительной разметки объекта. Для изображений с разметкой был выбран оптимальные детектор и дескриптор особых точек и пороговое значение, обеспечивающее отсутствие ложных срабатываний. После чего были вычислены глобальные координаты особых точек, на основе которых было проведено объединение частных трёхмерных моделей, полученных на первом шаге.

Для каждого этапа реконструкции были приведены соответствующие результаты.

Список литературы

1. Szeliski R. Computer Vision: Algorithms and Applications. 2010. P. 344–377.
2. Hartley R., Zisserman A. Multiple View Geometry in Computer Vision. 2003.
3. Karami E., Prasad S., Shehata M. Image matching using SIFT, SURF, BRIEF and ORB: Performance comparison for distorted images. 2017.
4. Rublee E., Rabaud V., Konolige K., Bradski G. ORB: an efficient alternative to SIFT or SURF // Computer Vision (ICCV). 2011. P. 2564–2571.
5. Leutenegger S., Chli M., Siegwart R. BRISK: Binary Robust Invariant Scalable Keypoints // Computer Vision (ICCV). 2011. P. 2548–2555.
6. Patent SIFT [Электронный ресурс]: <https://www.google.com/patents/US6711293> (дата обращения: 14.05.18)
7. Patent SURF [Электронный ресурс]: <http://www.google.com/patents/US20090238460> (дата обращения: 14.05.18)
8. Calonder M., Lepetit V., Strecha C., Fua P. BRIEF: Binary Robust Independent Elementary Features // 11th European Conference on Computer Vision (ECCV). 2010. P. 778–792.
9. Huang J., Cowan B. Simple 3D Reconstruction of Single Indoor Image with Perspective Cues // Canadian Conference on Computer and Robot Vision. 2009.

10. Töppe E., Oswald M., Cremers D., Rother C. Oswald M. R. et al. Non-parametric single view reconstruction of curved objects using convex optimization // Joint Pattern Recognition Symposium. 2009. P. 171–180.
11. Töppe E., Oswald M., Cremers D., Rother C. Image-based 3D Modeling via Cheeger Sets. // Asian Conference on Computer Vision. 2010. P. 53–64.
12. Töppe E., Oswald M., Cremers D., Rother C. Silhouette-Based Variational Methods for Single View Reconstruction. // Video Processing and Computational Video. 2011. P. 104–123.
13. Töppe E., Oswald M., Cremers D. Fast and globally optimal single view reconstruction of curved objects // Computer Vision and Pattern Recognition (CVPR). 2012. P. 534-541.
14. Malik J., Rosenholtz R. Computing local surface orientation and shape from texture for curved surfaces // International journal of computer vision. 1997. V. 23. №. 2. P. 149–168.
15. Ikeuchi K., Horn B. Numerical shape from shading and occluding boundaries // Artificial intelligence. 1981. V 17. №. 1–3. P. 141–184.
16. Nagai T., Naruse T., Ikehara M., Kurematsu A. Hmm-based surface reconstruction from single images // Image Processing. 2002. V. 2. P. II-II.
17. Tanskanen P. Live metric 3d reconstruction on mobile phones. // Computer Vision (ICCV). 2013. P. 65–72.
18. Carltona C., Mitchellb S., Lewis P. Preliminary application of Structure from Motion and GIS to document decomposition and taphonomic processes // Forensic science international. 2018. V. 282. P. 41–45.

19. Micheletti N., Chandler J. H., Lane S. N. Structure from motion (SFM) photogrammetry. 2015.
20. Pock T., Cremers D., Bischof H., Chambolle A. An algorithm for minimizing the Mumford-Shah functional // IEEE 12th International Conference on. 2009. P. 1133–1140.
21. Rodehorst V., Koschan A. Comparison and Evaluation of Feature Point Detectors. // 5th International Symposium Turkish-German Joint Geodetic Days. 2006.
22. FLANN - Fast Library for Approximate Nearest Neighbors [Электронный ресурс]: http://www.cs.ubc.ca/research/flann/uploads/FLANN/flann_manual-1.8.4.pdf (дата обращения: 14.05.18).
23. Правила подготовки модели к печати [Электронный ресурс]:<https://habrahabr.ru/post/196182/>(дата обращения: 14.05.18).