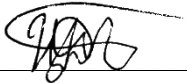


МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО
ОБРАЗОВАНИЯ
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ТЕХНОЛОГИЧЕСКИЙ
УНИВЕРСИТЕТ «МИСиС»

ИНСТИТУТ _____ ИТАСУ
КАФЕДРА _____ ИНЖЕНЕРНОЙ КИБЕРНЕТИКИ
НАПРАВЛЕНИЕ _____ 01.03.04 ПРИКЛАДНАЯ МАТЕМАТИКА

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА БАКАЛАВРА

на тему: Алгоритм декомпозиции человеческой позы на элементы простой формы по фотоснимку.

Студент _____  _____ Д. А. Цой
Руководитель работы _____ Р. В. Сенченко

Работа рассмотрена кафедрой и допущена к защите в ГЭК

Заведующий кафедрой _____ О. А. Ускова
Директор института _____ С. В. Солодов

Москва 2020

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО
ОБРАЗОВАНИЯ

«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ТЕХНОЛОГИЧЕСКИЙ
УНИВЕРСИТЕТ «МИСиС»

УТВЕРЖДАЮ

Институт ИТАСУ

Кафедра ИНЖЕНЕРНОЙ КИБЕРНЕТИКИ

Зав. Кафедрой Ускова О. А.

Направление 01.03.04 Прикладная математика

«26» декабря 2019 г.

**ЗАДАНИЕ
НА ВЫПОЛНЕНИЕ ВЫПУСКНОЙ КВАЛИФИКАЦИОННОЙ
РАБОТЫ БАКАЛАВРА**

Студенту группы БПМ-16-2, Цой Дарье Александровне

1. Тема работы: Алгоритм декомпозиции человеческой позы на элементы простой формы по фотоснимку
2. Цель работы: Создание математического, алгоритмического и соответствующего программного обеспечения, позволяющего моделировать трехмерную фигуру позы человека, разложенную на элементарные геометрические тела.
3. Исходные данные: Фотографии, содержащие людей в полный рост
4. Основная литература, в том числе:
 - 4.1. Монография, учебники и т.п.: Васильев Ф. П. Методы оптимизации: учеб. пособие // Васильев Ф.П., Потапов М.М., Будак Б.А., Артемьева Л.А. — М.: изд-во Юрайт, 2019
 - 4.2. Отчеты по НИР, диссертации, дипломные работы и т.п. _____
 - 4.3. Периодическая литература: Bogo F. Keep it SMPL: automatic estimation of 3d human pose and shape from a single image. // Bogo F., Kanazawa A., Lassner C., Gehler P. V., Romero J., and Black M. J. — In: ECCV, 2016
 - 4.4. Справочники и методическая литература (в том числе литература по методам обработки экспериментальных данных): Методические указания по оформлению выпускных квалификационных работ бакалавров

5. Перечень основных этапов исследования и форма промежуточной отчетности по каждому этапу:

5.1. Литературный обзор предметной области

5.2. Формулировка содержательной и математической постановок задачи

5.3. Сбор необходимых исходных данных

5.4. Разработка алгоритмов

5.5. Разработка программного обеспечения

5.6. Тестирование разработанного алгоритма

5.7. Анализ полученных результатов

5.8. Подготовка текста выпускной квалификационной работы, доклада и презентации

6. Аппаратура и методики, которые должны быть использованы в работе: Объектно-ориентированный подход, алгоритмы вычислительной геометрии и методы оптимизации

7. Использование ЭВМ: Разработка программного обеспечения на языке программирования Python с использованием библиотек NumPy, SMPL, PIL, VTKPlotter, Pyrender.

8. Перечень (примерный) основных вопросов, которые должны быть рассмотрены и проанализированы в литературном обзоре: Методы восстановления позы человека в трехмерном пространстве по фотоснимку, алгоритмы вычислительной геометрии

9. Перечень (примерный) графического и иллюстрированного материала актуальность поставленной задачи, описание предметной области, содержательная постановка задачи, математическая постановка задачи, функциональная схема, схема алгоритмов, основные экранные формы, полученные результаты и их анализ, выводы

10. Руководитель работы _____ ст.преп. Сенченко Р.В.

Дата выдачи задания 26.12.2019

Задание принял к исполнению студент _____ 

РЕФЕРАТ

Выпускная квалификационная работа выполнена на 50 страницах, содержит 46 рисунок, 12 формул, список использованных источников из 35 пунктов.

Ключевые слова: ОБРАБОТКА ИЗОБРАЖЕНИЙ, ГРАДИЕНТНЫЙ СПУСК, ДЕТЕКТИРОВАНИЕ ЧЕЛОВЕКА, ВОССТАНОВЛЕНИЕ ПОЗЫ, ГЕОМЕТРАЛЬНЫЙ МЕТОД РИСОВАНИЯ, ГЕСТАЛЬТА

Выпускная квалификационная работа посвящена разработке системы, предназначенной для восстановления трёхмерной позы человека по двумерному изображению и упрощению формы объекта геометрическими фигурами.

Целью данной работы является разработка математического, алгоритмического и программного обеспечения, позволяющего моделировать трехмерную фигуру позы человека, разложенную на элементарные геометрические тела. Объект на выходе является аналогом гестальты – шарнирной деревянной куклы для рисования.

В ходе работы проанализированы методы и подходы для распознавания людей на изображениях и восстановления их трехмерной фигуры. Кроме того, рассмотрены методы вычислительной геометрии для построения выпуклых оболочек.

Основным результатом данной работы являются методы с сопутствующим программным обеспечением для отображения обобщенной формы позы человека поверх исходного изображения, а также сохранения трехмерного объекта для экспортирования в трехмерные редакторы. Данная программа протестирована на выборке фотографий с различными позами, ракурсами, телосложением и количеством людей.

Разработанное программное обеспечение рассчитано на использование в образовательных целях в художественных учреждениях, а также для самостоятельного обучения людей, желающих развить пространственное мышление и навыки рисования.

В данный момент на рынке имеются программные продукты, которые также используют трехмерную графику для отображения человеческих фигур. Однако отличительной особенностью предложенной системы является возможность восстановления позы по изображению.

Программа реализована на языке программирования Python с использованием библиотек NumPy, SMPL, PIL, VTKPlotter, Pyrender.

ABSTRACT

The final qualifying paper contains 50 pages, 46 figures, 12 formulas, a list of references that includes 35 titles.

Keywords: IMAGE PROCESSING, GRADIENT DESCENT, HUMAN DETECTION, POSE ESTIMATION, GEOMETRIC METHOD IN DRAWING, GESTALTA

The final qualifying paper is devoted to the development of a system designed to restore a three-dimensional human pose from a two-dimensional image and simplify the shape of an object with geometric figures.

The purpose of this work is to develop a mathematical and algorithmic software allows to simulate a three-dimensional figure of a person's pose, decomposed into elementary geometric bodies. The output object is an analogue of a gestalta - an articulated wooden artist's dummy.

During the work, methods and approaches for Human recognition from images and reconstruction of 3D body shape are analyzed. In addition, computational geometry methods for constructing convex hulls are considered.

The main result of this work is methods with related software for displaying a generalized form of a person's pose over the original image, as well as saving a three-dimensional object for export to three-dimensional editors. This program was tested on samples of photos with various poses, foreshortening, physique and number of people.

The developed software is designed for educational use in art institutions, as well as for self-learning of people who want to develop spatial thinking and drawing skills.

There are currently software products on the market that also use three-dimensional graphics to display human figures. However, one of the features of the proposed system is the ability to restore the pose from the image.

The program is implemented in the Python programming language using the libraries NumPy, SMPL, PIL, VTKPlotter, Pyrender.

СОДЕРЖАНИЕ

СПИСОК ИСПОЛЬЗУЕМЫХ ОСНОВНЫХ СОКРАЩЕНИЙ	7
ВВЕДЕНИЕ	8
1 АНАЛИТИЧЕСКИЙ ОБЗОР ЛИТЕРАТУРЫ.....	10
1.1 Основные понятия нейронных сетей.....	10
1.2 Методы вычислительной геометрии	11
1.2.1 Построение минимальных выпуклых оболочек	11
1.2.2 Алгоритм Грехэма	11
1.2.3 Теорема о разделяющей оси.....	12
1.2.4 Обнаружение столкновений.....	14
1.2.5 Алгоритм построения ОБВ	15
1.2.6 Основные понятия компьютерной графики	16
1.3 Детектирование частей тела человека.....	16
1.3.1 DeepPose	17
1.3.2 Forward-or-Backward Information (FBI)	20
1.3.3 SMPL.....	20
2 СПЕЦИАЛЬНАЯ ЧАСТЬ	24
2.1 Содержательная постановка задачи.....	24
2.2 Математическая постановка задачи.....	25
2.3 Адаптация математического аппарата к входным данным	27
2.3.1 Первый этап: выбор базовой модели.....	27
2.3.2 Второй этап: разбиение входных данных на подзадачи.....	28
2.3.3 Третий этап: выбор метода построения фигур	29
2.4 Идея алгоритма.....	32
2.5 Описание алгоритма	34
2.6 Функциональная схема	39
2.7 Тестирование и результаты	40
ЗАКЛЮЧЕНИЕ.....	47
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ.....	48

СПИСОК ИСПОЛЬЗУЕМЫХ ОСНОВНЫХ СОКРАЩЕНИЙ

ПО – программное обеспечение

МВО – минимальная выпуклая оболочка

ААВВ – ограничивающий параллелепипед, параллельный координатным осям

ОВВ – ориентированный ограничивающий параллелепипед

FBI – Forward-or-Backward Information

SCAPE – Shape Completion and Animation for PEople

SMPL – A skinned multi-person linear model

ВВЕДЕНИЕ

Данная работа находится на стыке искусства и науки. Связь между рисунком и математикой заметили еще в эпоху Возрождения. Большой вклад в этой области внес Леонардо да Винчи, а также Дюрер, Альберти, Микеланджело, Ашбе и множество других выдающихся личностей [1].

Так, к примеру, итальянский учёный, математик и художник Альберти считал рисование такой же точной, подвергаемой законам и правилам наукой, как и математика [2]. Он заложил научное обоснование в обучение изобразительному искусству, заключающемуся не в механическом копировании видимого образа, а в форме упражнений, способных развивать мышление учеников. Как точно потом сформулировал эту мысль Микеланджело "Рисуют головой, а не руками". Также большой вклад в теорию рисования внес великий Леонардо да Винчи.

Человеческий мозг устроен так, что его мышление, память и другие формы познания действительности строятся на базовом восприятии образов [3]. Следовательно, большое значение на раннем этапе развития человека играет пространственное воображение.

Помимо перечисленных аспектов, образное мышление является важным навыком для математиков [4]. Для решения сложных задач необходимо уметь оперировать абстрактными понятиями в математическом пространстве, что довольно сложно сделать без представления общей структуры. И, естественно, данное умение должно быть хорошо развито у художников. У многих начинающих рисовальщиков возникает сложность в отражении правильных пропорций и перспективных искажений объектов сложной формы [5]. Особенно часто проблема возникает с изображением человека. Так в конце 19 века выделились две методики в школьном преподавании рисунка, а именно геометральный и натуральный методы [6].

Первый основывается на обобщении сложной формы до простых геометрических фигур: параллелепипедов, сфер, призм, конусов и т.п. Такое представление помогает легче улавливать правильные расстояния, размеры и пропорции. После построения такой конструкции набросок уточняется и детализируется, как бы переходя от простого к сложному.

Второй метод предполагает перенос видимых линий и кривых непосредственно, опираясь только на интуитивное восприятие.

Одним из основоположников геометрального метода является Ашбе [7]. В его школу-студию в Мюнхене съезжались многие ученики из всех стран Европы, в том числе

и из России. Также сторонниками данного принципа были Дюрер [8], Чистяков [9], Сапожников [10] и многие другие. Примечательно то, что с 1888 года геометральный метод был введен во всех общеобразовательных школах. В настоящее время он часто встречается в программах художественных школ и различных пособиях по рисованию.

В данной работе автором рассмотрено использование методов машинного обучения для решения задачи восстановления фигуры и позы человека в пространстве по двумерному изображению. Целью данного исследования является разработка алгоритма декомпозиции 3D модели человеческой фигуры на простые геометрические тела, точно повторяющие силуэт позы.

Прототипом трехмерного объекта на выходе является Гестальта (от нем. Gestalt – образ, форма) – небольшая шарнирная кукла с гнущимися конечностями (рисунок 1). В ее создании был заложен геометральный принцип, так как формы в ней максимально упрощены, что дает возможность изучать расположение теней и световых бликов на достаточно простом объекте.



Рисунок 1 – Гестальта

Такой болванчик подходит для изучения правильных пропорций человека и тренировки его рисования с разных углов обзора. Но в силу своей механической природы, эти фигурки ограничены в вариативности возможных представлений поз.

Чтобы уйти от ограниченности физической модели было предложено смоделировать данную концепцию посредством компьютерных технологий. Такой инструмент сможет частично заменить наставления педагогов в области изобразительного искусства или расширить спектр их возможностей для преподавания.

1 АНАЛИТИЧЕСКИЙ ОБЗОР ЛИТЕРАТУРЫ

1.1 Основные понятия нейронных сетей

Со временем электронные машины обогнали по скорости вычислений человеческие возможности. Как следствие, люди стали задумываться о возможности создания искусственного интеллекта, совмещающего сознательное мышление человека и мощность расчетов компьютера. Так, для понимания принципов работы мозга, исследователи рассмотрели биологическую составляющую работы нервной системы. При попытке воссоздания мозговых процессов появились искусственные нейронные сети (ИНС) [11].

Нейронная сеть представляет собой совокупность функций-нейронов, соединенных в слои и связанных связями – синапсами – с помощью сумматоров. В основе концепции модели лежат процессы работы человеческого мозга. Когда на вход поступают сигналы, активируется первый слой и продолжает распространять «импульсы» дальше по сети. Каждый активный нейрон взвешивает входные значения, суммирует их и при достаточной силе возбуждает следующий слой.

Для того, чтобы достичь желаемый результат на выходе модели, необходимо настроить определенным способом веса нейронов. Эта процедура получила название «процесс обучения» [12].

Сверточная нейронная сеть (CNN) – архитектура нейронной сети, разработанная для решения задач, связанных с изображениями [13]. Основной характерной чертой является наличие операции свертки [14], при которой каждый фрагмент изображения умножается на фильтр, а результат записывается в новую матрицу меньшего размера (рисунок 1.1).

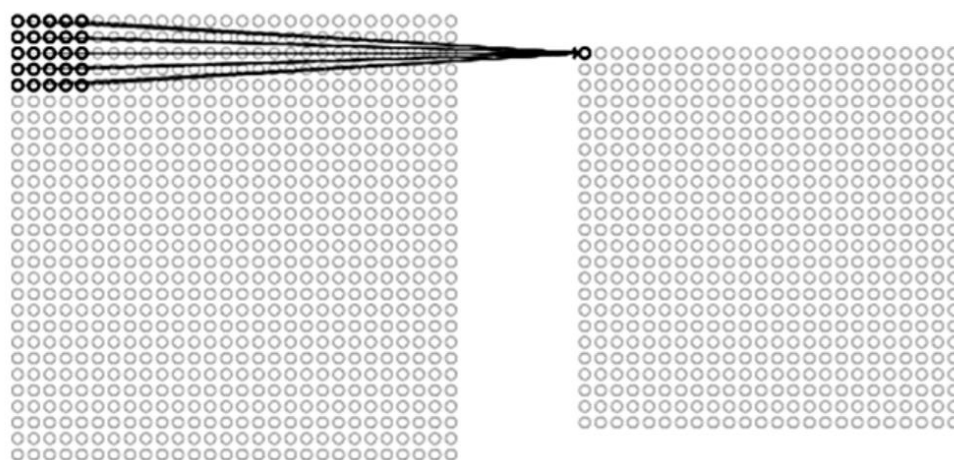


Рисунок 1.1– Пример операции свертки фильтром размера 5 x 5

1.2 Методы вычислительной геометрии

1.2.1 Построение минимальных выпуклых оболочек

Пусть задано конечное множество точек A в пространстве \mathbb{R}^2 . Оболочкой этого множества называется любая замкнутая кривая без петель, содержащая все исходные элементы A (рисунок 1.2). Кривая называется выпуклой, если ни одна касательная не пересекает ее ни в одной точке. Тогда соответствующая ей оболочка также считается выпуклой.

Минимальной выпуклой оболочкой (МВО) называют самую короткую выпуклую кривую [15].

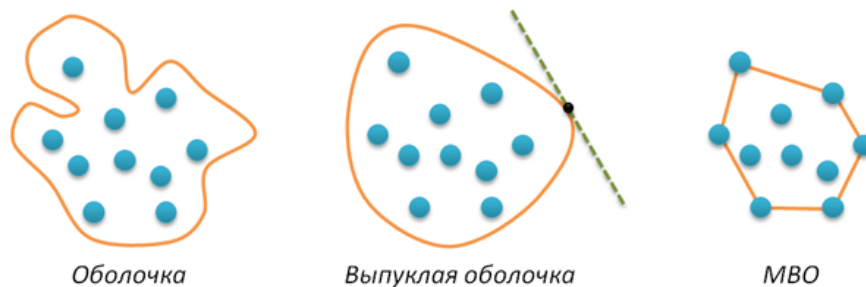


Рисунок 1.2 – Иллюстрация оболочек конечного множества

ВПО активно используется в задачах вычислительной геометрии. Также они имеют ряд практических применений в области поиска границ и распознавании образов.

Задача. Задано множество точек A . Построить их выпуклую оболочку.

Для данной задачи было разработано много различных подходов с разной временной сложностью [16-18]. Но на практике чаще всего используется алгоритм Грэхема, опубликованный в 1972 году [19]. Сложность этого метода – линейно-логарифмическая.

1.2.2 Алгоритм Грэхема

На первом шаге алгоритма выбирается самая левая нижняя точка p_1 множества S , что выполняется за время $O(n)$. Далее от каждой текущей точки строится луч и следующей точкой берется первая по часовой стрелке. Таким образом, выполняется сортировка по возрастанию углов (рисунок 1.3).

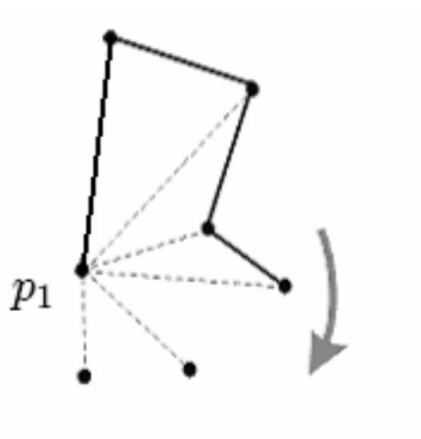


Рисунок 1.3 – Порядок обхода точек в алгоритме Грэхема

Для хранения отсортированного массива используют двусвязный список (p_1, p_2, \dots, p_n) . При его прохождении ищутся точки, входящие в ВПО, а остальные удаляются.

Далее будем считать, что три точки a , b и c образуют *левый поворот*, если точка c находится слева от прямой, проходящей через a и b .

Для построения выпуклой оболочки на каждом шаге берется тройка точек из списка, проверяется ориентация поворота и, если он является левым, средняя из точек удаляется. На ее место встает следующий элемент списка. Процесс останавливается, когда на проверку снова приходит первая точка множества p_t . Данная схема получила название «метода трех монет».

Так как каждая точка в списке рассматривается только один раз, алгоритм выполняет не более n шагов. В итоге сложность алгоритма Грэхема складывается из процесса сортировки и обхождения списка (1):

$$O(n \log n) + O(n) = O(n \log n). \quad (1)$$

1.2.3 Теорема о разделяющей оси

Чаще всего для определения пересечения двух тел (выпуклых множеств) используют теоремы о разделяющей оси (англ. Separating Axis Theorem) [20]. Она позволяет вычислить нормаль к рассекающей прямой и степень проникновения.

Теорема: два выпуклых объекта пересекаются тогда и только тогда, когда существует плоскость (для двумерного случая - прямая), такая, что одна геометрия лежит по одну ей сторону, а другая - по другую (рисунок 1.4).

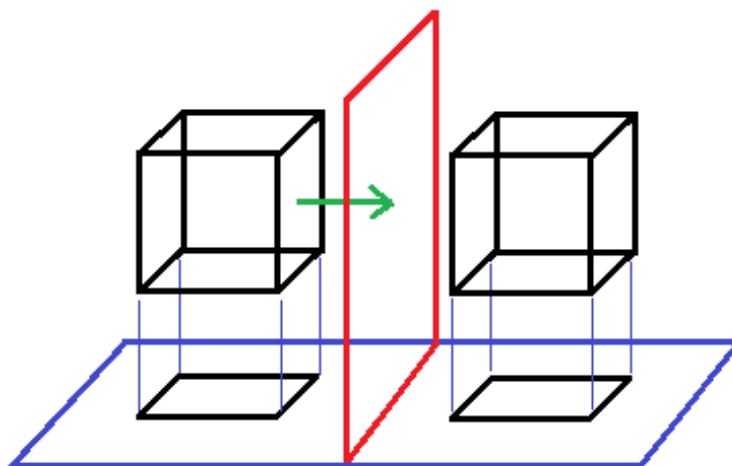


Рисунок 1.4 – Проверка на пересечение объектов по теореме о разделяющей оси

Следствие: если построить к разделяющей перпендикулярную плоскость, то проекции фигур на полученную плоскость тоже не будут пресекаться.

Замечание: разделяющих плоскостей (линий) может быть несколько.

Утверждение: разделяющая линия существует тогда и только тогда, когда существует разделяющая линия, параллельная ребру прямоугольника (рисунок 1.5).

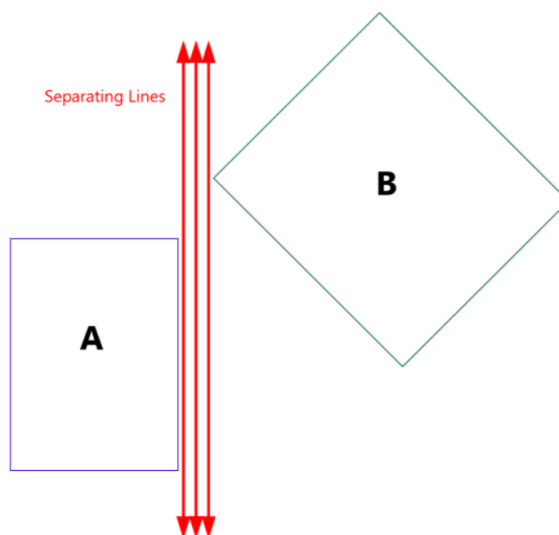


Рисунок 1.5 – Разделяющие линии для прямоугольников A и B

Другими словами, должна быть хотя бы одна разделяющая линия, параллельная ребру прямоугольника A или B, чтобы существовали другие разделяющие линии.

1.2.4 Обнаружение столкновений

Обнаружение столкновений (англ. Collision detection) – задача выявления пересечения двух тел. Данная проблема возникла в области компьютерной графики и игровых движков. Также активно применяется в робототехнике [21].

Главной ее целью является уменьшение вычислительных операций для обнаружения коллизий. Вместо того, чтобы обрабатывать сложные по структуре трехмерные объекты, было предложено оперировать более простыми примитивами, а именно: ограничивающими объемами (англ. Bounding volume). Так, с утерей некоторой точности, экономится огромное количество времени и ресурсов. Особенно когда необходимо обрабатывать большое число объектов в одной симуляции.

Проще говоря, ограничивающие объемы — это простые геометрические фигуры, куда вписываются более сложные по форме и структуре объекты.

Существует несколько способов представления объектов для обнаружения столкновений [22]. Но самыми часто используемыми являются ограничивающие параллелепипеды, параллельные координатным осям (англ. Axis-Aligned Bounding Boxes – AABB) и ориентированные ограничивающие параллелепипеды (англ. Oriented Bounding Box – OBB) (рисунок 1.6).

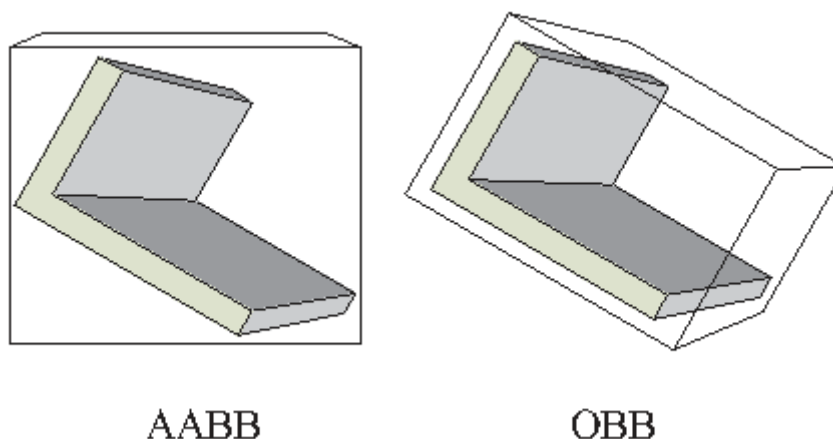


Рисунок 1.6 – Разница между AABB и OBB

Разница между ними в том, что AABB задается ребрами, параллельными осям, а OBB не зависит от мировой системы координат. В первом случае получается очень грубое приближение, но оно обусловлена элементарными операциями для обнаружения пересечений. То есть, если скорость важнее точности, то этот метод является самым простым и легким для решения данной задачи.

Преимуществом ОВВ является постоянство размеров при вращении тела в пространстве. Ребра строятся относительно локальной системы координат, привязанной к самому объекту. То есть, он более приближен к габаритам аппроксимируемого тела. Однако вычисления коллизий становятся сложнее и дольше.

1.2.5 Алгоритм построения ОВВ

Пусть дано множество точек A . Ставится задача найти минимальный ориентированный ограничивающий прямоугольник.

Первым этапом будет построение минимальной выпуклой оболочки (см. раздел 1.2.1). Она является кусочно-линейной замкнутой кривой, такой что все точки из A расположены либо на ней самой, либо во внутренней области (рисунок 1.7).

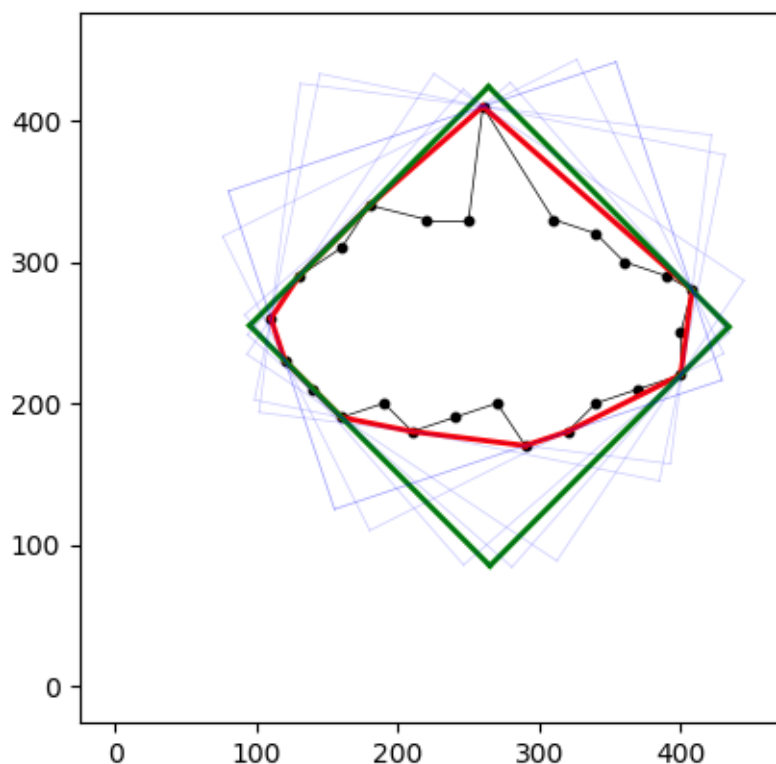


Рисунок 1.7 – Этапы работы алгоритма ОВВ

Из постановки задачи ясно, что за пределами искомого прямоугольника не должно находиться не одной точки. Это можно трактовать так, что контур ОВВ не должен пересекаться с МВО данного множества [23].

Тогда из утверждения теоремы о разделяющей оси (см. раздел 1.2.4) ребра минимального прямоугольника будут параллельны одному из отрезков минимальной выпуклой оболочке данного множества.

Следовательно, вторым этапом необходимо вычислить углы наклонов для каждого отрезка МВО и рассчитать объемы ААВВ для всех найденных направлений.

Минимальный из полученных прямоугольников будет являться решением задачи (см. на Рисунке 1.7 прямоугольник с жирным контуром).

1.2.6 Основные понятия компьютерной графики

В компьютерной графике трёхмерные объекты описываются полигональной сетки (англ. polygon mesh), составленной из массивов вершин, рёбер и граней, которые задают форму объекта [24].

Рендеринг или отрисовка (англ. rendering – «визуализация») – это процесс программной генерации изображения высокого качества по моделям из компьютерной графики.

Здесь модель — это комплексное описание трехмерного объекта, в терминах геометрических примитивов, положения камеры, информации об освещении, текстуре поверхности и т.п.

Скининг (англ. Skinning, skin – кожа) – это процесс привязки полигональной сетки к скелету. Он позволяет моделировать движение сложных моделей персонажей задавая только движение каркаса фигуры.

1.3 Детектирование частей тела человека

«Определение позы» (pose estimation) – это метод компьютерного зрения, который детектирует фигуры людей на изображениях и оценивает местоположение основных суставов [24].

Задача обнаружения человека получила широкий спектр применения в различных областях, а именно в: системах безопасности, в дополненной реальности, в робототехнике, для идентификации личности и др.

Данная задача имеет два пути решения:

- алгоритмы, использующие 2D изображение;
- алгоритмы, использующие стереоизображения для формирования пространственной модели сцены.

Хотя и использование нескольких источников изображений, например принцип работы Kinect камеры, дают более точные результаты, в данной работе основной целью является рассмотрение алгоритмов, принимающих на вход именно двумерные данные.

В свою очередь, оценка по одному снимку может классифицироваться на 2D и 3D предсказания позы (рисунок 1.8), т.е. разметку суставов тела на изображении (в терминах значений пикселей) или расположение узловых точек в трехмерном пространстве.

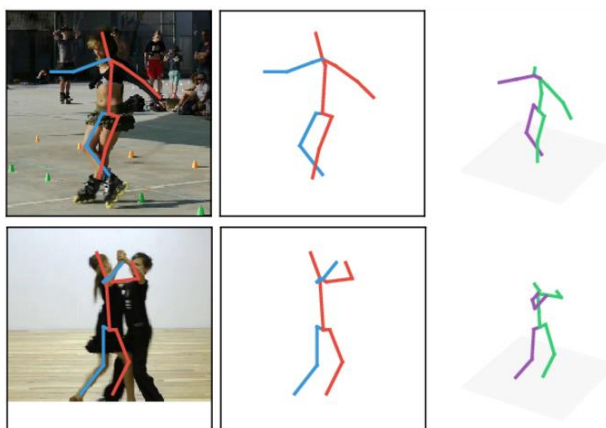


Рисунок 1.8 – Результаты алгоритма FBI

1.3.1 DeepPose

Для начала рассмотрим простой алгоритм оценки позы в двумерном пространстве. DeepPose [25] был предложен Александром Тошевым и Кристианом Сегеди на международной конференции по компьютерному зрению CVPR еще в 2014 году и отражает базовые принципы подходов к решению данной задачи.

Данный метод предлагает представлять скелет человека в виде графа, где вершинами будут суставы, а ребрами — кости. А дальше ставится задача построить математическую модель, которая будет описывать вероятность появления того или иного сустава в конкретном месте изображения, а также будет учитывать, насколько реалистично эти суставы расположены по отношению друг к другу.

Данный алгоритм использует глубокие нейронные сети и представляет собой задачу регрессии [26]. То есть, для каждого сустава на фотографии нужно определить его координаты.

Обозначим входное изображение как x , а вектор позы как

$$y = (y_1, \dots, y_i, \dots), \quad i \in \{1, \dots, k\},$$

$$N(y_i, b) = \begin{pmatrix} 1 & 0 \\ \frac{1}{b_w} & 0 \\ 0 & \frac{1}{b_h} \end{pmatrix} (y_i - b_c). \quad (2)$$

За $N(x, b)$ обозначим часть исходной картинке x в рамке b . Все исходное изображение обозначим за b_0 .

Введем функцию нормализации вектора позы $\psi(x, \theta) \in R^{2k}$. Тогда в исходных координатах он примет вид (3):

$$y = N^{-1}(\psi(N(x), \theta); b_0), \quad (3)$$

где θ — параметр модели.

В качестве функции ψ выступает нейронная сеть, веса которой описываются θ [27]. То есть мы подаем на вход трехканальное изображение фиксированного размера, а получаем на выходе k пар координат суставов (рисунок 1.9):

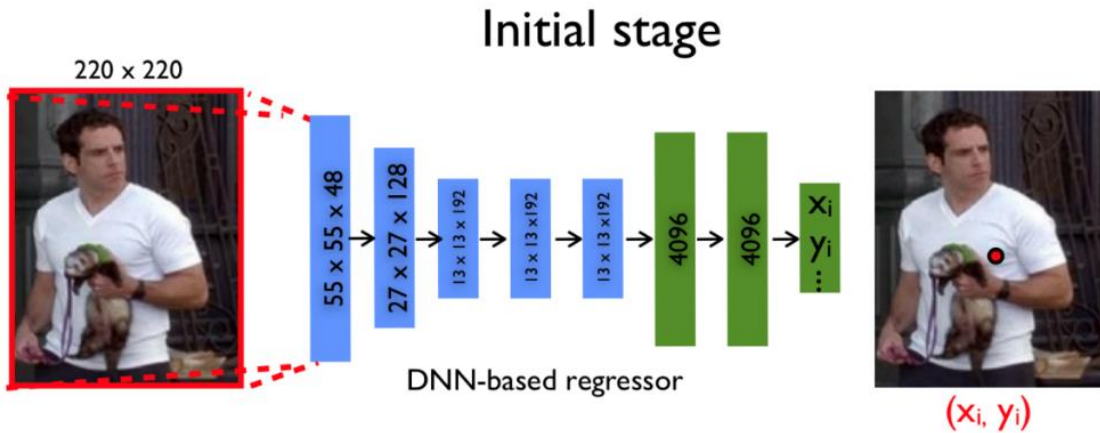


Рисунок 1.9 – Схема нейронной сети DeepPose

Для обучения регрессии найдем такие параметры модели, чтобы минимизировать сумму квадратов отклонений предсказанных координат от истинных. То есть идеальные параметры (4):

$$\sum_{(x,y) \in D_N} \sum_{i=1}^k \|y_i - \psi(x, \theta)\|_2^2. \quad (4)$$

Таким образом, мы пытаемся восстановить исходные координаты суставов независимо друг от друга.

Первым шагом $s = 1$ рассматривается тривиальная область b^0 , в которой предсказываются предварительные координаты суставов (5):

$$\text{Stage 1: } y^1 \leftarrow N^{-1}(\psi(N(x; b^0), \theta_1); b^0) \quad (5)$$

Далее, на каждой новой итерации ($s = 2 \dots S$) и для каждого сустава $i \in \{1, \dots, k\}$ координаты y_i^{s-1} уточняются с помощью модели $\psi(x, \theta_s)$ (рисунок 1.10) (6-7):

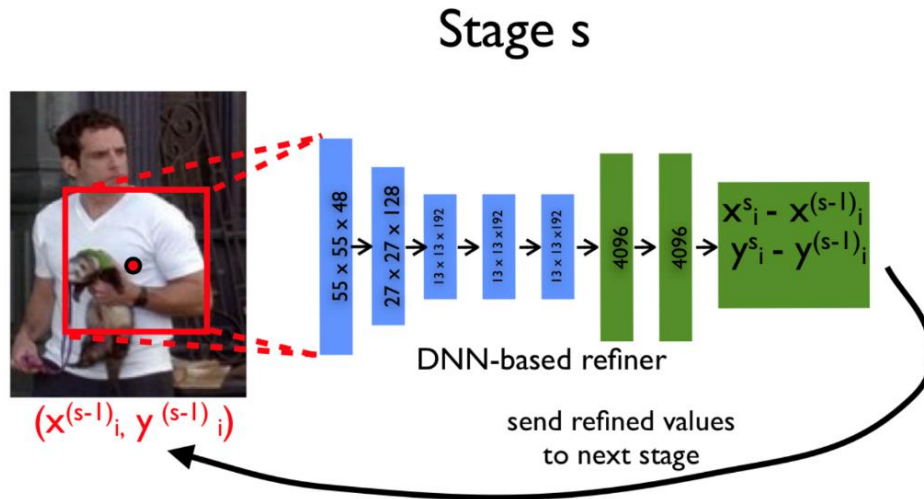


Рисунок 1.10 – Шаг метода ограничивающей области

$$\text{Stage } s: y_i^s \leftarrow y_i^{s-1} + N^{-1}(\psi_i(N(x; b_i^{s-1}), \theta_s); b_i^{s-1}), \quad (6)$$

$$b_i^s \leftarrow (y_i^s, \sigma \text{ diam}(y^s), \sigma \text{ diam}(y^s)). \quad (7)$$

Для оценки качества разметки, полученной от обучаемой сети, авторы метода предлагают использовать две следующие метрики:

- процент корректно распознанных частей тела (Percentage of Correct Parts, PCP);
- количество корректно предсказанных суставов (Percent of Detected Joints, PDJ).

1.3.2 Forward-or-Backward Information (FBI)

Метод прямой или обратной информации (Forward-or-Backward Information, FBI) предсказывает не только координаты суставов на входном изображении, но и рассчитывает скелет позы в трехмерном пространстве. Авторы алгоритма предлагают использовать сведения об удалённости каждой части человеческого тела от камеры (рисунок 1.11).

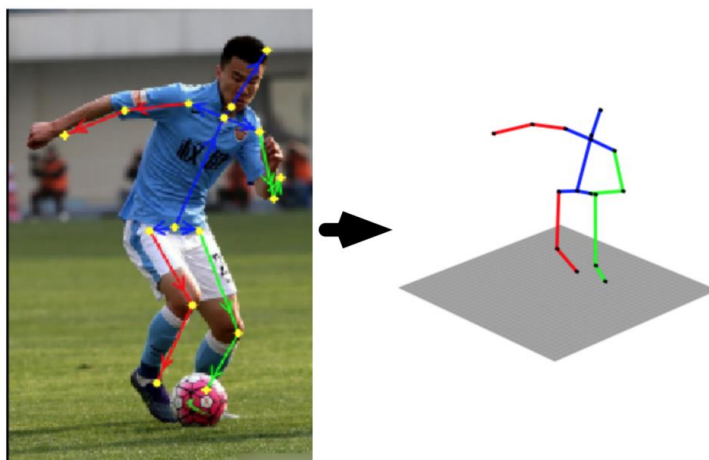


Рисунок 1.11 – Определение позы человека с помощью информации о дальности расположения его конечностей относительно камеры

Для решения задачи используется две сверточных нейронных сети: для детектирования координат суставов тела на изображении и для предсказания глубины точки относительно камеры [28].

Преимущество данного метода в том, что координата в двухмерном пространстве имеет несколько образов в трехмерной системе. То есть, различные позы могут иметь одинаковые проекции на плоскость изображения. Для точного предсказания положения частей тела необходимо больше информации, как минимум расстояние по третьей оси, перпендикулярной плоскости камеры.

1.3.3 SMPL

Трехмерные объекты кодируются множеством вершин и заданием на них полигональных сеток (см. раздел 1.2.7). Чем сложнее модель, тем большим количеством

вершин она кодируется. Однако увеличение размерности точек поверхности тела ведет к увеличению вычислений при рендеринге.

Одной из проблем компьютерной графики является обеспечение эффективности использования вычислительных ресурсов. Авторы статьи *Keep it SMPL: automatic estimation of 3d human pose and shape from a single image* [29] заметили, что для таких объектов, как человек, трехмерное полигональное представление может быть сжато до пространства меньшего размера. Так как все люди имеют сходное строение, предлагается брать заранее разработанную базовую сетку поверхности тела и задавать набор параметров, а именно: высоту, окружность торса, размер живота, позу и т. п. Это представление имеет гораздо меньший объем и требует меньше вычислительных расчетов.

SMPL – это статистическая модель, которая кодирует людей с двумя типами параметров:

– Параметр формы: вектор из 10 скалярных значений, каждое из которых можно интерпретировать как величину расширения / сжатия человеческого объекта в некотором направлении. Например, на рисунке 1.12 продемонстрировано изменение параметра высоты и веса.

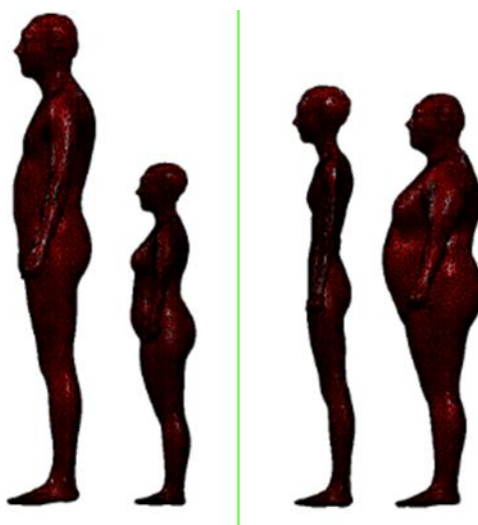


Рисунок 1.12 – Визуализация изменения двух параметров формы SSMPL

– Параметр позы: вектор относительных поворотов размера 23×3 , для пространственного расположения различных частей тела (рисунок 1.12).

Pose: Rotation of 23 joints

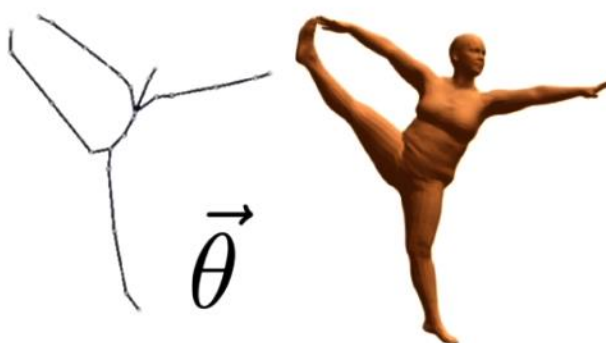


Рисунок 1.12 – Повороты, формирующие позу модели человека

В данной работе был разработан алгоритм автоматической оценки трехмерной позы человеческого тела, а также его трехмерной формы по одному двумерному изображению. Было показано, что полигональную сетку модели человека можно эффективно сформировать основываясь на информации о расположении суставов, т. е. точек пространства, отражающих ключевые точки позы, такие как голова, кисть, локоть, колено и т. п. (рисунок 1.13) Эта проблема является сложной из-за сложности человеческого тела, артикуляции, окклюзии, одежды, освещения и присущей неоднозначности в отображении в 3D из 2D.

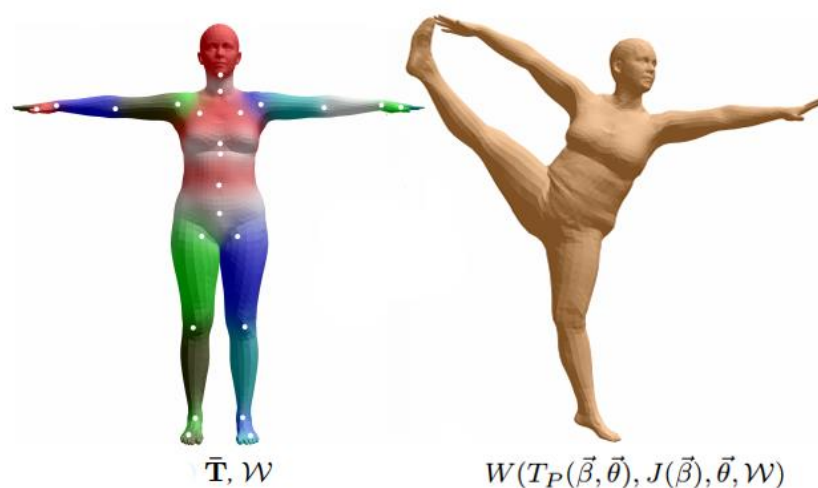


Рисунок 1.13 – Формирование позы и формы тела под влиянием точек суставов

Чтобы решить эту проблему, авторы сначала используем метод на основе сверточных нейронных сетей, DeepCut [30], для прогнозирования расположения суставов на двумерном изображении. Затем они подгоняют статистическую модель формы тела

SMPL [31] к 2D суставам. Это делается путем минимизации целевой функции, которая штрафует ошибку между проецируемыми суставами 3D-модели и обнаруженными 2D-суставами.

Существует еще одна широко известная статическая модель SCAPE [32], в которой используется метод на основе треугольной деформации [33]. Но в отличие от нее, SMPL использует скиннинг на основе вершин [34]. То есть, форма одной части тела представляется как вектор конкатенированных смещений вершин начальной сетки с $N = 6890$ точками и $K = 23$ соединениями.

Тестирование обоих алгоритмов показывает, что SMPL значительно выигрывает в производительности. Это показано на рисунке 1.14 где по вертикали отложено время в секундах на кадр, необходимое для генерации вершин, а по горизонтали их количество.

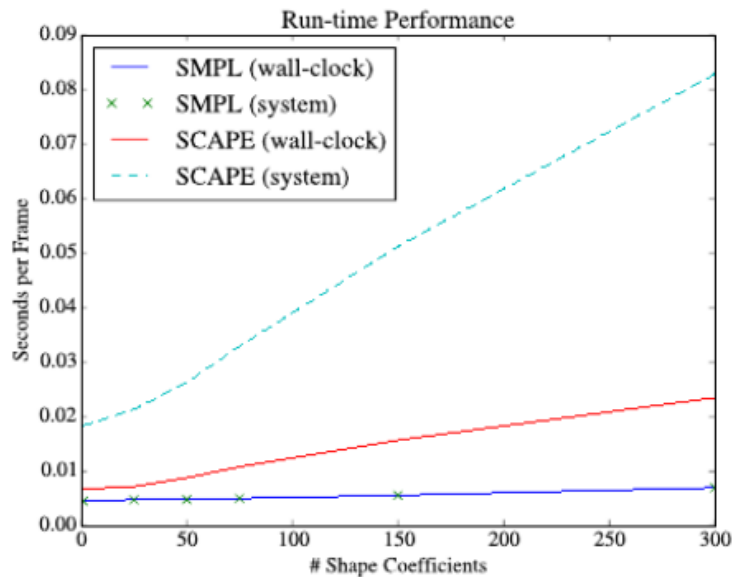


Рисунок 1.14 – Тестирование производительности моделей SMPL и SCAPE

2 СПЕЦИАЛЬНАЯ ЧАСТЬ

2.1 Содержательная постановка задачи

Целью выпускной квалификационной работы является создание математического, алгоритмического и соответствующего программного обеспечения, позволяющего моделировать трехмерную фигуру позы человека, разложенную на элементарные геометрические тела, а именно: параллелепипеды (рисунок 2.1).

Для решение поставленной проблемы необходимо выполнить следующие пункты:

1) Провести анализ существующих подходов к задаче распознавания и восстановления позы и фигуры человека по двумерному изображению.

2) Выбрать подходящий метод, эффективный по времени обработки фотографии.

3) Выбрать количество многогранников для построения модели, передающих общий силуэт тела человека.

4) Разработать математическо-алгоритмическое обеспечение для декомпозиции полученного объекта на простые элементы, аппроксимирующие первоначальную конструкцию.

5) Провести анализ и визуальную оценки полученного результата и сделать выводы.

Разработанный алгоритм должен быть инвариантен к количеству точек объекта, для которого строится параллелепипед. А модель на выходе должна учитывать телосложение индивида и пропорции каждой части тела.

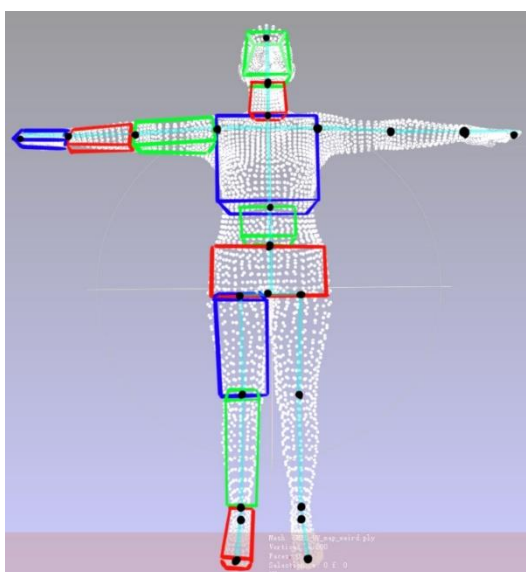


Рисунок 2.1 – Схематичное изображение искомой модели

2.2 Математическая постановка задачи

Пусть $P = \{p_1, p_2, \dots, p_N\}$ – множество N точек из \mathbb{R}^3 , образующие фигуру человека в трехмерном пространстве.

На множестве P задано разбиение P_1, P_2, \dots, P_K ($P_i \cap P_j = \emptyset$ при $i \neq j$ и $\bigcup_{i=1}^K P_i = P$), где P_i – множество точек одной из K частей тела (например, рука или голова).

Позы человека задается в виде кусочно-линейной кинематической модели, которая описывает объект как объединение нескольких шарнирно закрепленных твердых тел. Эту структуру можно представить в виде K соединенных отрезков (рисунок 2.2). Обозначим их через $O = \{o_1, o_2, \dots, o_K\}$, $o \in \mathbb{R}^{2 \times 3}$. Таким образом, каждая ось задается парой точек трехмерного пространства.

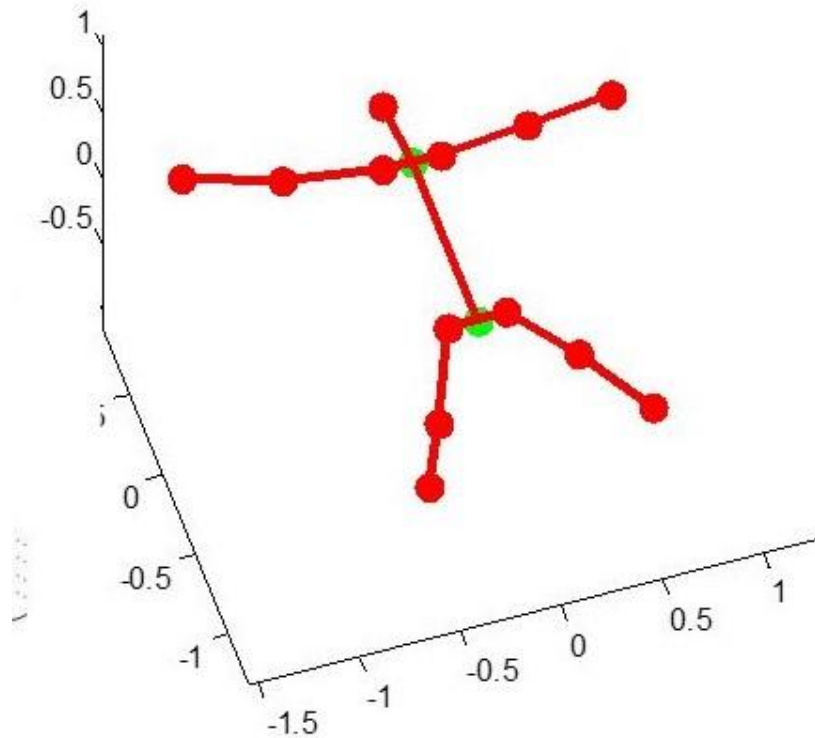


Рисунок 2.2 – Кинематическая модель тела человека

Пусть $R = \{r_1, r_2, \dots, r_K\}$ – множество K параллелепипедов, центры оснований которых совпадают с концами соответствующих отрезков из O . Т. е. задана биекция $\chi: O \rightarrow R$.

Каждая фигура $r \in \mathcal{R}$ характеризуется набором параметров $l(r_i): \mathcal{R} \rightarrow \mathbb{R}^3, i = \overline{1, K}$, а именно: $l_i(r) = \{\varphi, h, w\}$ – угол поворота вокруг оси o_i , высота и ширина оснований соответственно.

Пусть $M = \{m_1, m_2, \dots, m_K\}$ – множество K шестиугольных призм, основания которых имеют равные углы под 60° (см. Рисунок 2.3). Т. е. задана биекция $\psi: R \rightarrow M$.

Каждая фигура $m \in \mathfrak{R}$ также имеет набор параметров $u_i(m): \mathfrak{R} \rightarrow \mathfrak{R}^2, i = \overline{1, K}$. При этом они зависят от параметров прямоугольника $r: u_i(m) = q(l(r_i)) = \{h_m, w_m\}$.

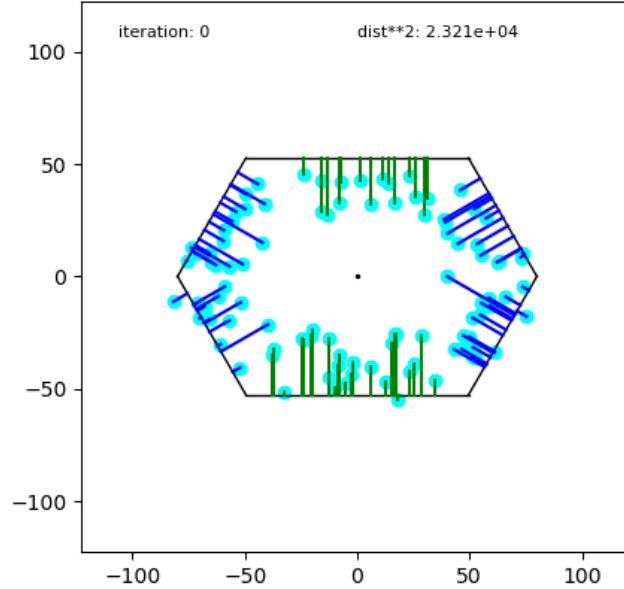


Рисунок 2.3 – Расстояния от точек до сторон основания призмы из M

Очевидно, что каждая фигура m имеет 8 сторон. Введем расстояние $\rho'(p, m)$ как евклидово расстояние от точки p до ближайшей грани призмы m (рисунок 2.4).

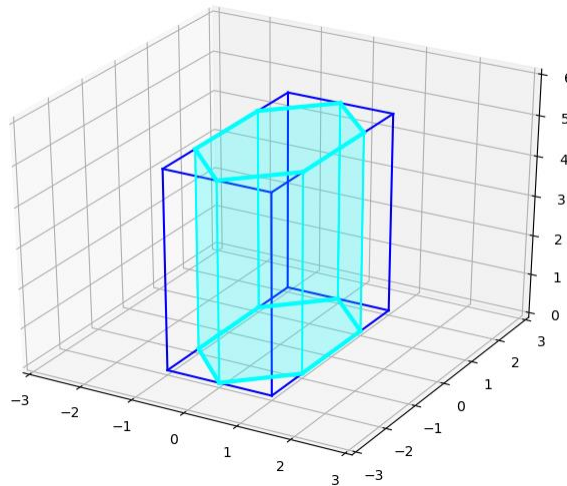


Рисунок 2.4 – Искомый параллелепипед и вписанная призма в пространстве

Ставится задача для $1 < i < N$ найти такие параметры $l(r(P_i, m_i))$, что сумма расстояний от всех точек P_i до фигуры m_i была бы минимальной (2.1), т. е.

$$Q(P_i, m_i) = \sum_{t=1}^{|P_i|} \rho'(p_t, m_i)^2 \rightarrow \min_{l(r(P_i, m_i))}, \quad \text{для } \forall p_t \in P_i. \quad (8)$$

2.3 Адаптация математического аппарата к входным данным

Содержательная постановка была разбита на три блока, которые отражают постепенный ход решения поставленной задачи. Сначала идет подготовительный этап, в котором определяется входное множество. Далее идет разбиение входных данных на мелкие подзадачи. И, наконец, разрабатывается итоговый алгоритм. Т. е. проблема решается в три шага:

- 1) Выбор метода для распознавания человека на изображении.
- 2) Декомпозиция полученных входных данных на $K = 17$ подмножеств для каждой части тела.
- 3) Ход решения подзадачи для одного многогранника из общей конструкции искомой модели.

2.3.1 Первый этап: выбор базовой модели

Для основы разработки алгоритма был взят метод SMPL, описанный в разделе 1.3.3 (рисунок 2.5). Эта модель не только предсказывает позу, но и восстанавливает характерные особенности фигуры человека, такие как рост, телосложение и т. п., что является ключевым фактором выбора для решения поставленной задачи. А также методика в основе имеет превосходство по производительности и точности среди других подобных моделей.

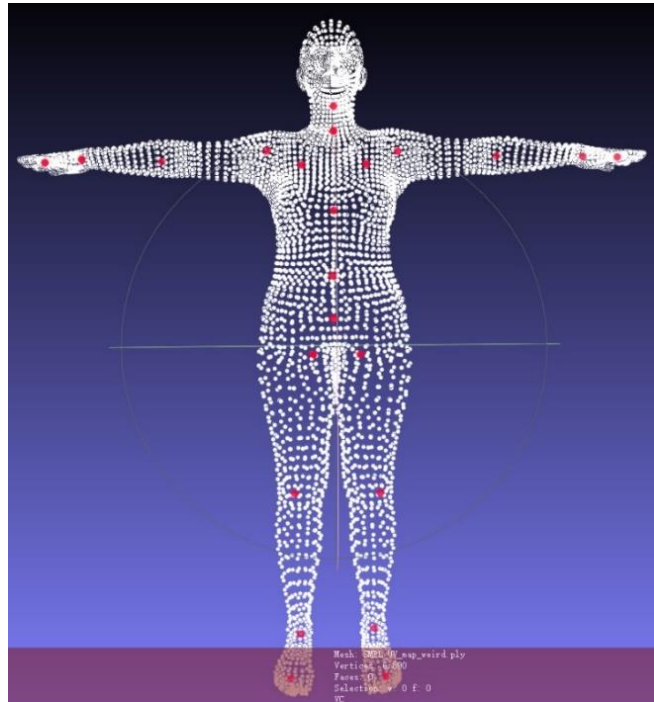


Рисунок 2.5 – Выходные данные SMPL алгоритма

2.3.2 Второй этап: разбиение входных данных на подзадачи

Рассматриваемое программное обеспечение получает на вход двумерное изображение человека и пропускает его через сверточную нейронную сеть (рисунок 2.6). На выходе рассчитывается матрица размера 6890×3 , которая представляет собой координаты (x, y, z) 6890 точек, образующих силуэт фигуры человека (рисунок 2.6) и 23×3 точки размеченных суставов позы.

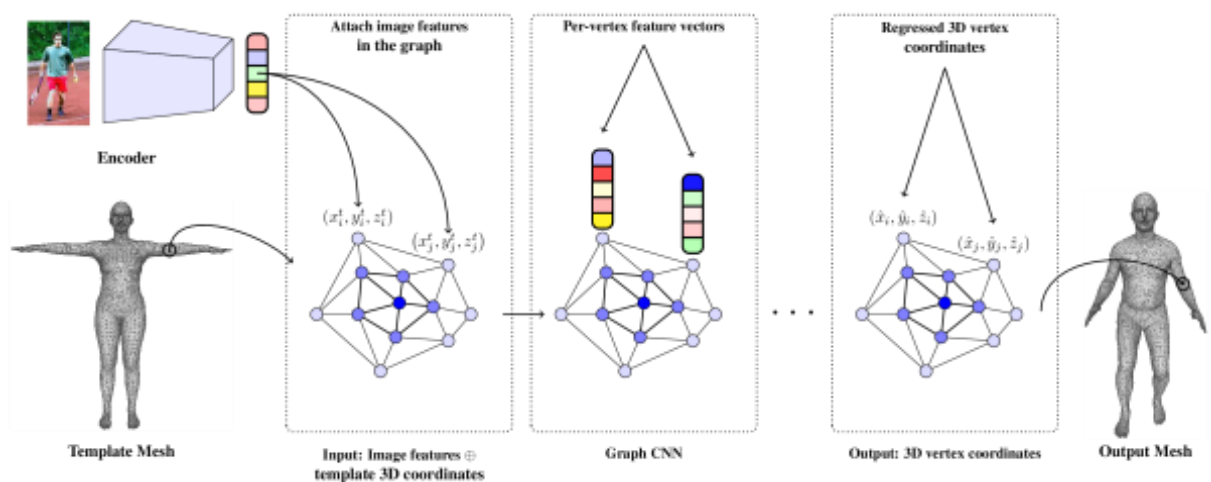


Рисунок 2.6 – Этапы работы SMPL

Задача о наибольшем пустом прямоугольнике

В текущей работе поверхности тел далеки от идеальной формы параллелепипеда, а наибольший пустой прямоугольник, по определению, не содержит в себе другие объекты. В связи с этими фактами, объем рассчитываемой фигуры будет сильно меньше объема аппроксимируемого тела (рисунок 2.8). Как следствие, первоначальный силуэт тела человека будет сильно искажен в итоговой модели. Поэтому данный метод плохо решает поставленную задачу.

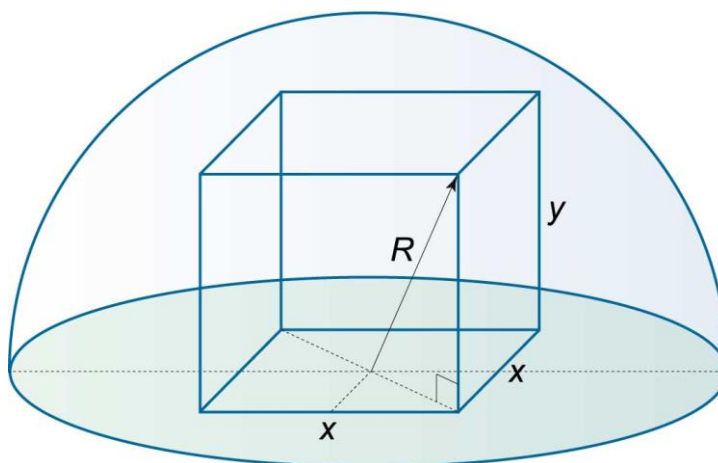


Рисунок 2.8 – Решение задачи о максимальном пустом параллелепипеде

АВВВ

Ограничивающие параллелепипеды часто используются в физических движках различных игр для обнаружения столкновений или пересечений различных объектов между собой. Множество таких областей называется хитбоксами или «зонами поражений».

В отличие от предыдущего подхода, такой контейнер, наоборот, должен содержать в себе все множество точек. Но в решаемой задаче его объем будет все также сильно отличаться от первоначального пространственного тела. Из-за этого фигура будет выглядеть «раздутой» (рисунок 2.9).

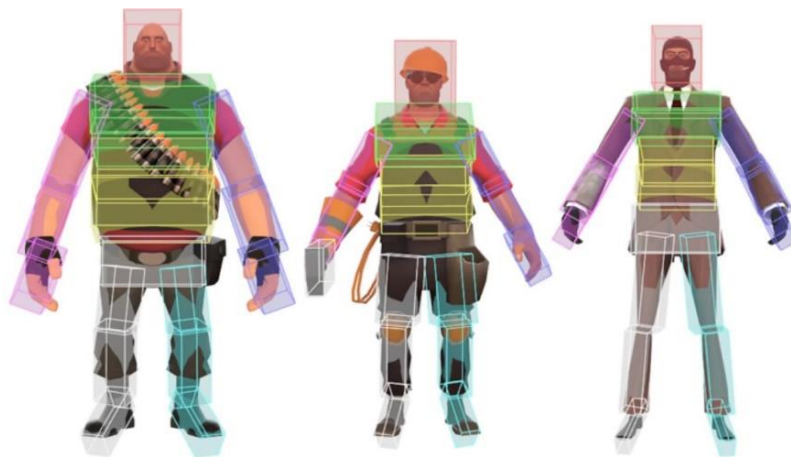


Рисунок 2.9 – Хитбоксы

ОВВ

ОВВ или ориентированный ограничивающий параллелепипед является модификацией AABB. Его преимущество в том, что он не привязан к координатным осям. Т. е. за счет вращений можно найти минимальную по объему фигуру.

Но методом экспериментов было выявлено, что изменение направления осей симметрии у оснований параллелепипедов хоть и уменьшает занимаемую площадь, но не отражает конструкции первоначального объекта. Например, на рисунке 2.10 с проекцией головы явно видно, что фигура вытянут вдоль оси ординат. Задача состоит в том, чтобы найти такой бокс, который будет отражать структуру черепа, т.е. должны быть грани со стороны ушей. А алгоритм ОВВ нашел прямоугольник, повернутый под 45° (рисунок 2.10). Если восстановить его до параллелограмма, получится ребро по середине лица. Таким образом этот алгоритм не решает проблему данной работы.

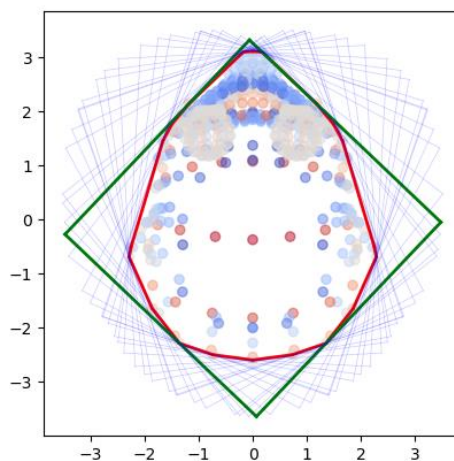


Рисунок 2.10 – Результат алгоритма ОВВ

Задач об окаймляющем параллелепипеде

В ходе выпускной квалификационной работы был разработан новый метод, который был назван «Задач об окаймляющем параллелепипеде».

За основу был взят тот факт, что природные тела имеют неровную поверхность и различные выступы. Чтобы оставить только главные черты некоторого объекта, можно вынести за область многогранника его незначительные, по сравнению с общим объемом, части. Например, при аппроксимации головы, можно оставить за границей параллелепипеда нос и уши. Тогда художник сможет уловить правильные пропорции элемента фигуры, а детали дорисовать потом.

Так как части тела в основном имеют цилиндрическую или сферическую формы, чтобы не сильно обрезать выступающие скругленные контуры, было принято решение увеличить число углов основания аппроксимирующей фигуры.

Таким образом, ставится задача найти такой равноугольный шестиугольник, что сумма расстояний от точек пространства до граней вписанной в него шестиугольной призмы была бы минимальной (см. раздел 2.2). Следовательно, необходимо решить задачу оптимизации.

2.4 Идея алгоритма

Рассмотрим некоторое подмножество P_i из P , например голову. Далее спроецируем все точки на плоскость, перпендикулярную отрезку o_i (рисунок 2.11).

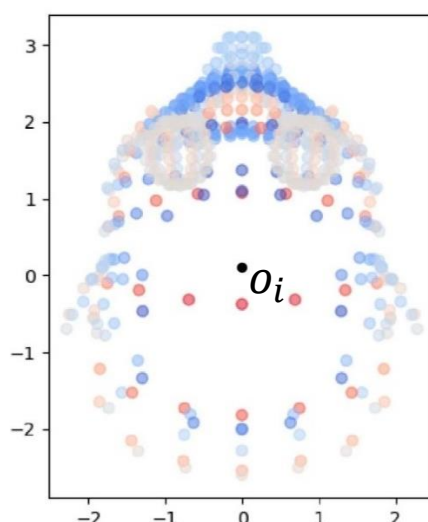


Рисунок 2.11 – Проекция точек головы на плоскость, перпендикулярную оси o_i

Тогда, чтобы построить параллелепипед r_i , удовлетворяющий заданным требованиям, достаточно найти основание в виде прямоугольника с центром в точке проекции o_i и сместить его вдоль этой оси. Т. е. необходимо вычислить параметры $l_i(r) = \{\varphi, h, w\}$.

Необходимо подобрать $l_i(r)$ так, чтобы аппроксимировать множество точек P_i ломанной, заданной сторонами прямоугольника.

По метод окаймляющего параллелепипеда для начала найдем аппроксимирующий шестиугольник (рисунок 2.12а), применим растяжение по одной из оси так, чтобы он стал правильным и найдем вписанную и описанную окружности (рисунок 2.12б). Далее найдем средний радиус (рисунок 2.12в) и сделаем обратное преобразование (рисунок 2.12г). Тогда описанный прямоугольник и есть необходимое основание r .

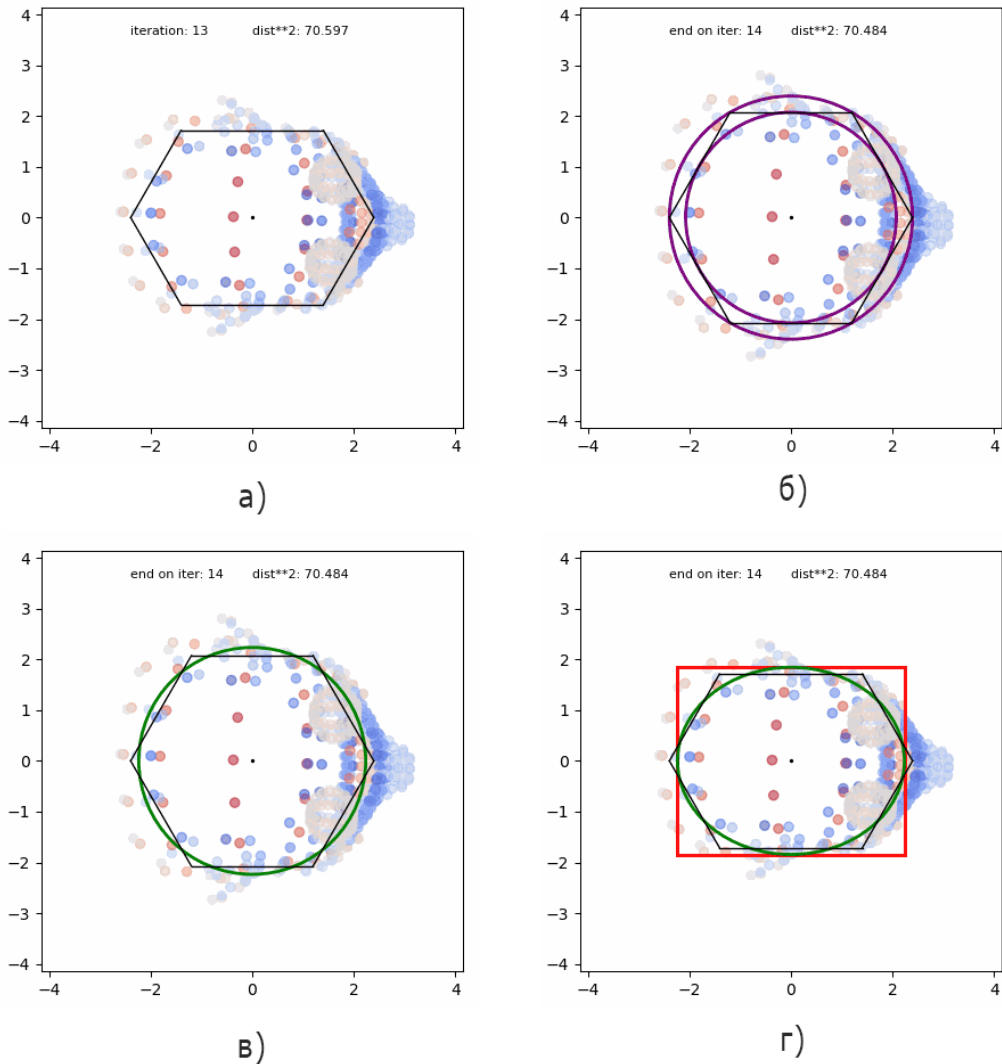


Рисунок 3.12 – Поиск основания прямоугольника

В идеальном случае, когда множество точек лежит на окружности, данным методом получаем описанный квадрат (рисунок 2.13).

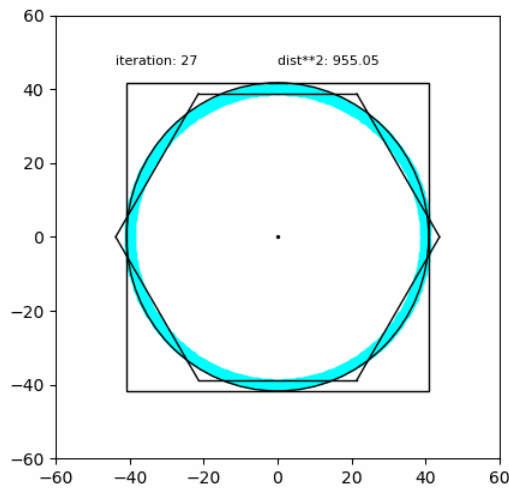


Рисунок 2.13 – Применение алгоритма к эталонному множеству

2.5 Описание алгоритма

Для простоты далее параметры шестиугольника $u(m) = \{h_m, w_m\}$ будем обозначать просто h, w .

Инициализируем первоначальную фигуру как минимальную рамку, содержащую все точки, т.е. $w = \max(X)$, $h = \max(Y)$, где X, Y – координаты точек по оси абсцисс и ординат соответственно. Если $h > w$, поворачиваем множество точек на 90° .

Проецируя точки фигуры на плоскость, мы не всегда можем быть уверенными в том, что ориентация контура совпадает с направлениями координатных осей (рисунок 2.14).

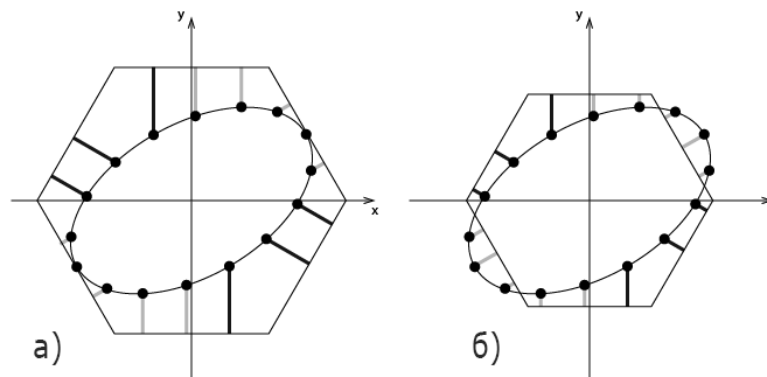


Рисунок 2.14 – Расстояния от вытянутого множества до сторон шестиугольника

Но, как можно заметить, мы имеем дело с приблизительно симметричными относительно диагоналей, объектами. Тогда сумма расстояний в I и III координатной четвертях и в II и IV должны быть приблизительно одинаковы (рисунок 2.13.а). При этом важно учитывать, где находится точка: внутри шестиугольника или снаружи.

Тогда, основываясь на разнице этих сумм, можно ввести эвристический алгоритм и итеративно поворачивать точки на заданный угол, пока разность сумм не станет меньше заранее заданной δ .

Пусть D_1^{in}, D_2^{in} – сумма расстояний от точек в I и III координатной четвертях, расположенных внутри шестиугольника. Аналогично D_1^{out}, D_2^{out} – суммы для точек вне фигуры. Тогда на каждом шаге цикла будем поворачивать фигуру на угол, равный

$$\theta_i = \frac{\pi}{4} \cdot k \cdot \left(\frac{D_1^{in} - D_2^{in}}{D_1^{in} + D_2^{in}} - \frac{D_1^{out} - D_2^{out}}{D_1^{out} + D_2^{out}} \right), \quad (9)$$

$$k \leq 1 \text{ при } D_1^{in} + D_2^{in} \neq 0 \text{ и } D_1^{out} + D_2^{out} \neq 0.$$

Сдвиг множества точек дает преимущество в обобщенном задании уравнений сторон шестиугольника. Тогда представим фигуру m как кусочно-заданную функцию (рисунок 2.15):

$$\left\{ \begin{array}{ll} y = h, & \text{при } tg60(x - w) + h \leq 0 \text{ и } -tg60(x + w) + h \geq 0 \\ y = -h, & \text{при } tg60(x - w) - h \geq 0 \text{ и } -tg60(x - w) - h \leq 0 \\ tg60(x - w) + y = 0, & \text{при } y \geq 0 \text{ и } y < h \\ tg60(x + w) + y = 0, & \text{при } y \leq 0 \text{ и } y > -h \\ -tg60(x - w) + y = 0, & \text{при } y \leq 0 \text{ и } y > -h \\ -tg60(x + w) + y = 0, & \text{при } y \geq 0 \text{ и } y < h \end{array} \right. \quad (10)$$

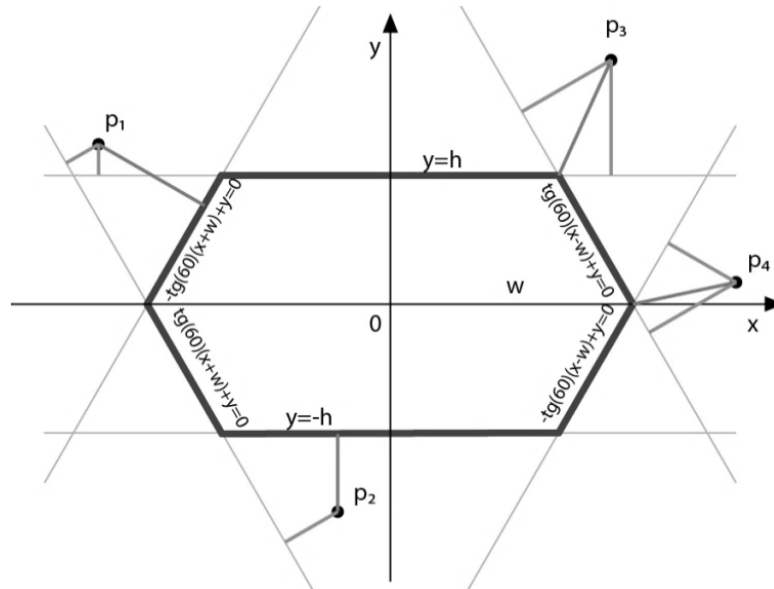


Рисунок 2.15 – Шестиугольник в виде кусочно-линейной функции

Расстояние ρ' от точки p до шестиугольника m было введено как евклидово расстояние до ближайшей из 6-ти сторон, но

– минимальное расстояние до прямой не всегда попадает на сторону шестиугольника, например точки p_1, p_2 ;

– есть области, из которых нельзя провести перпендикуляр к граням фигуры, например точки p_3, p_4 .

Поэтому добавим расчет длины отрезков от точки до углов многоугольника и определим области минимальных расстояний (рисунок 2.16).

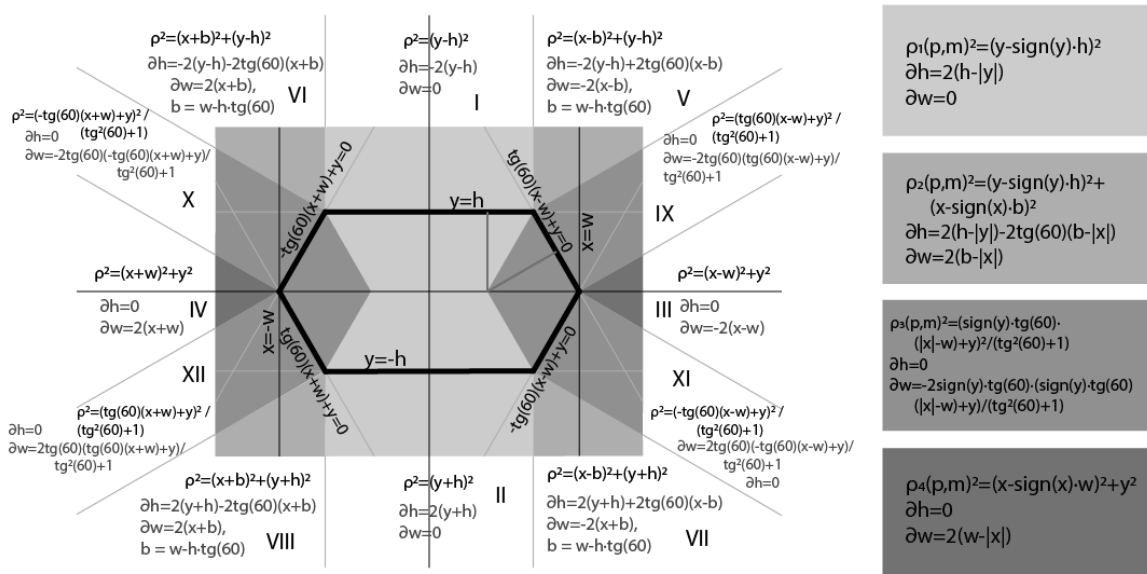


Рисунок 2.16 – Введение расстояний от точки до призмы m

Таким образом, имеет 12 разных секторов: 6, где расстояния считаются до сторон фигуры, и 6 – до углов. Производные по h и w рассчитываются по формулам области с минимальным расстоянием.

$$\rho'(p_t, m_i)^2 = \min_{z=1,4} (\rho_z(p_t, m_i)^2) =$$

$$= \min \left(\begin{array}{l} (y - \text{sign}(y) \cdot h)^2 \\ (y - \text{sign}(y) \cdot h)^2 + (x - \text{sign}(x) \cdot b)^2 \\ \frac{(\text{sign}(y) \cdot \text{tg}(60)(|x| - w + y))^2}{\text{tg}^2(60) + 1} \\ (x - \text{sign}(x) \cdot w)^2 + y^2 \end{array} \right), \quad b = w - \frac{h}{\text{tg}(60)} \quad (11)$$

$$dh = \begin{cases} 2(h - |y|), & \text{при } z = 1 \\ 2(h - |y|) - 2\text{tg}(60)(b - |x|), & \text{при } z = 2 \\ 0, & \text{иначе} \end{cases}$$

$$dw = \begin{cases} 0, & \text{при } z = 1 \\ 2(b - |x|), & \text{при } z = 2 \\ \frac{2\text{sign}(y) \cdot \text{tg}(60) \cdot (\text{sign}(y) \cdot \text{tg}(60)(|x| - w + y))}{\text{tg}^2(60) + 1}, & \text{при } z = 3 \\ 2(w - |x|), & \text{при } z = 4 \end{cases} \quad (12)$$

где z – номер индекса минимальной области из (2.2).

Функция минимизации в формуле 8 имеет 2 зависимые переменные h и w . Тогда решение можно найти методом градиентного спуска.

Алгоритм:

Инициализация: выбор параметра шага λ , параметр поворота k и параметры критерия останова поворота и градиентного спуска δ и ε соответственно.

Вход: множество точек $P = \{(x_i, y_i)\}_{i=1}^N$

Выход: пара (h, w)

Повторять:

$$h^{[k+1]} = h^{[k]} - \lambda \frac{\partial Q}{\partial h} Q(h^{[k]}, w^{[k]})$$

$$w^{[k+1]} = w^{[k]} - \lambda \frac{\partial Q}{\partial w} Q(h^{[k]}, w^{[k]})$$

$$\theta^{[k+1]} = \frac{\pi}{4} \cdot k \cdot \left(\frac{D_1^{in} - D_2^{in}}{D_1^{in} + D_2^{in}} - \frac{D_1^{out} - D_2^{out}}{D_1^{out} + D_2^{out}} \right),$$

при $D_1^{in} + D_2^{in} \neq 0$ и $D_1^{out} + D_2^{out} \neq 0$

Пока:

$$|w^{[k+1]} - w^{[k]}| > \varepsilon \text{ or } |h^{[k+1]} - h^{[k]}| > \varepsilon \text{ or } |\theta^{[k+1]}| < \delta.$$

2.6 Функциональная схема

Далее приведем функциональную схему всех этапов алгоритма (рисунок 2.17):

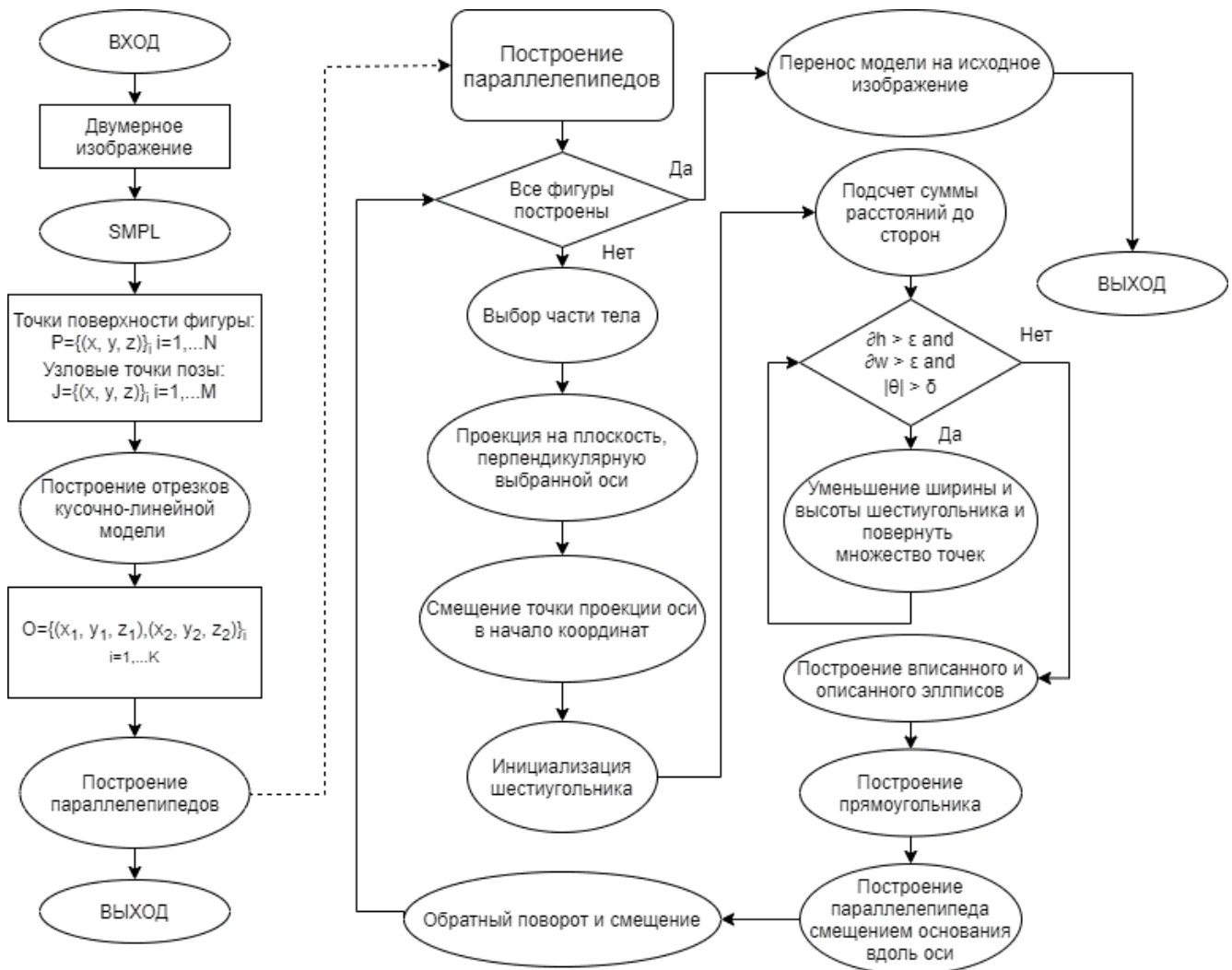


Рисунок 2.17 – Функциональная схема алгоритма

2.7 Тестирование и результаты

Для демонстрации алгоритма в качестве входных данных возьмем трехмерную модель головы человек (рисунок 2.18).

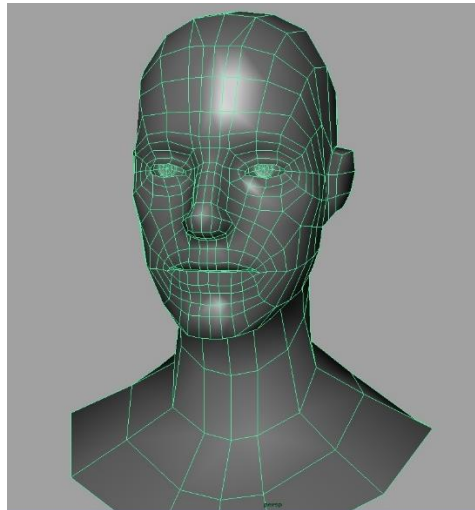


Рисунок 2.18 – Трехмерная модель головы человека

С помощью библиотеки `matplotlib` для `python` визуализируем полученный результат (рисунки 2.19-20):

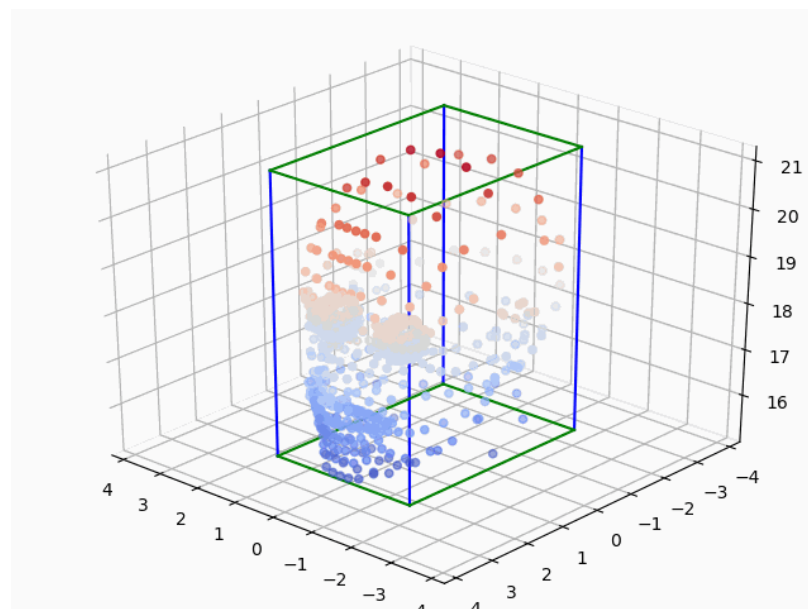


Рисунок 2.19 – Результат алгоритма окаймляющего параллелепипеда в изометрии

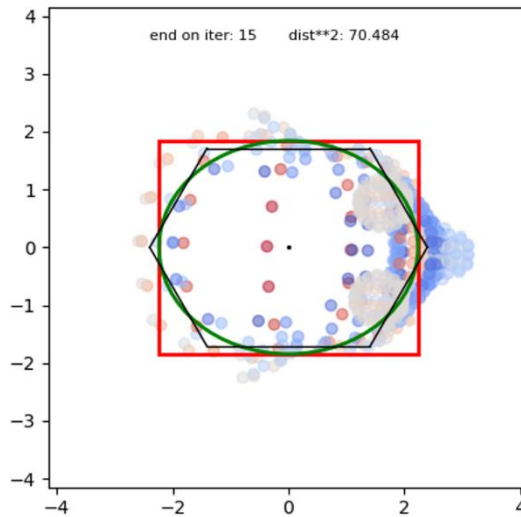


Рисунок 2.20 – Результат алгоритма окаймляющего параллелепипеда, вид сверху

На виде сбоку и спереди (рисунки 2.21-22) можно ясно увидеть, что только череп попал в многогранник, а нос и уши вышли за границу области. Таким образом, требование, определенное в разделе 2.4, было выполнено.

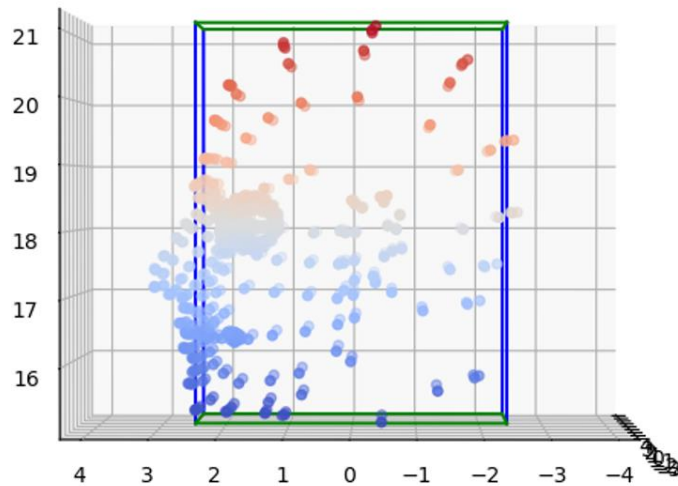


Рисунок 2.21 – Результат алгоритма окаймляющего параллелепипеда, вид сбоку

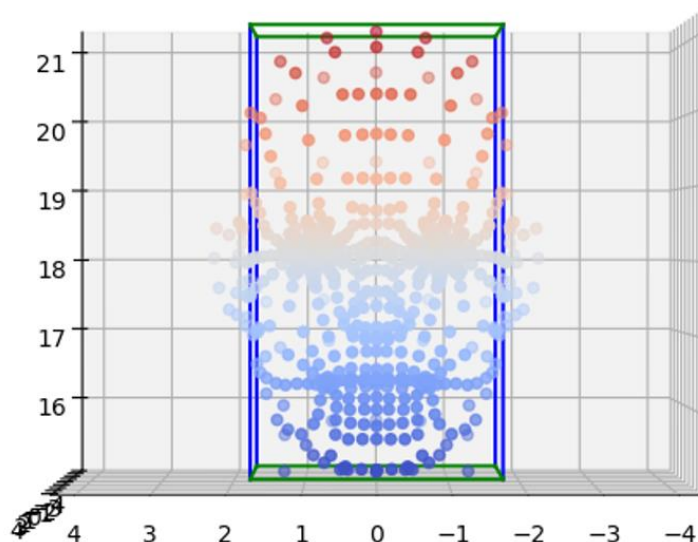


Рисунок 2.22 – Результат алгоритма окаймляющего параллелепипеда, вид спереди

Рассмотрим некоторые иллюстрации из учебных пособий по рисованию (рисунки 2.23-24). Как можно заметить, авторы учебных материалов предлагают на первом этапе построить область, ограничивающую только основной объем объекта, например, череп. Это позволит точно уловить нужные размеры и пропорции. А мелкие детали вроде ушей или носа оставить вне куба. Их художник наметит после основных элементов.

Сравнивая результаты программы с рекомендациями учебников, можно сделать вывод, что метод подходит для решения поставленной задачи.

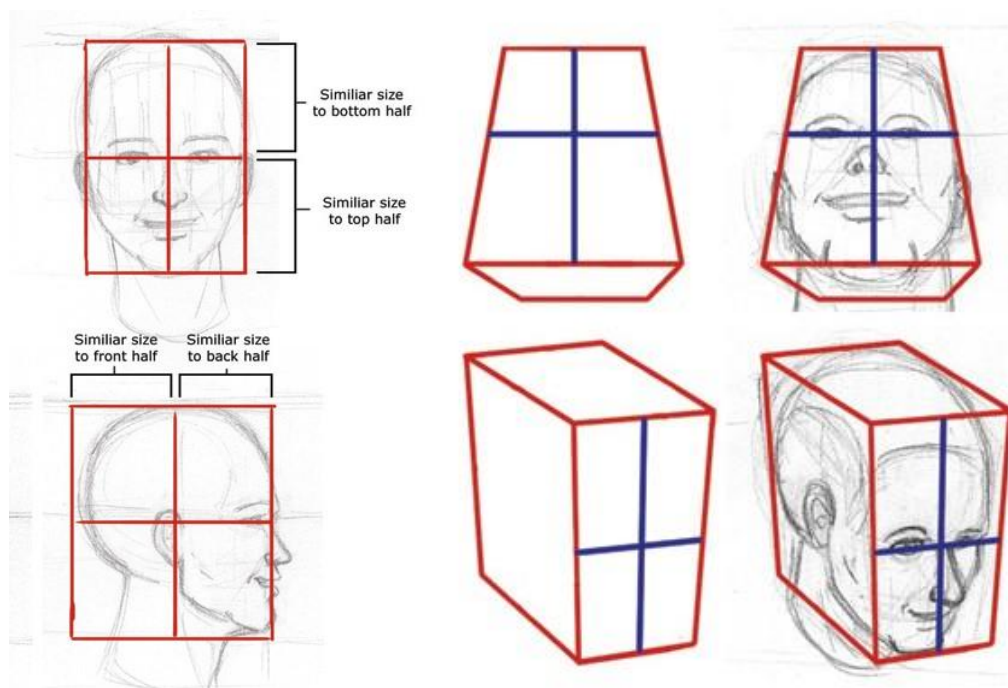


Рисунок 2.23 – Рекомендации к упрощению формы головы из пособия по рисованию

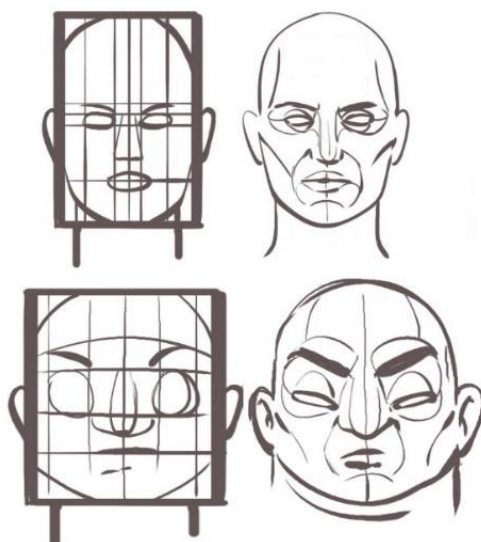


Рисунок 2.24 – Рекомендации к упрощению формы головы из пособия по рисованию

Теперь протестируем алгоритм на изображениях людей. Во-первых, программное обеспечение показало отличные результаты в нестандартных ракурсах и позах, например, на изображениях со спортивными упражнениями (рисунки 2.25-2.26).



Рисунок 2.25 – Результат работы программы

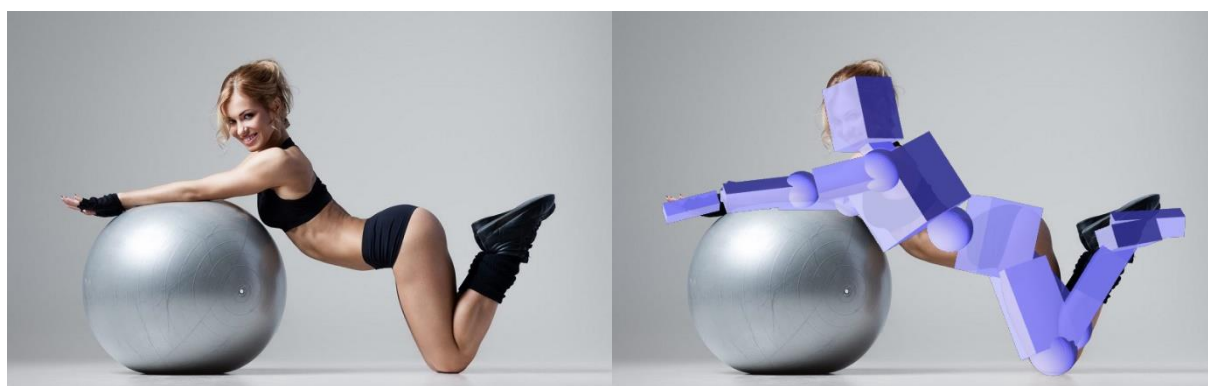


Рисунок 2.26 – Результат работы программы

Во-вторых, ПО позволяет восстанавливать несколько человек на одной фотографии (рисунок 2.27).

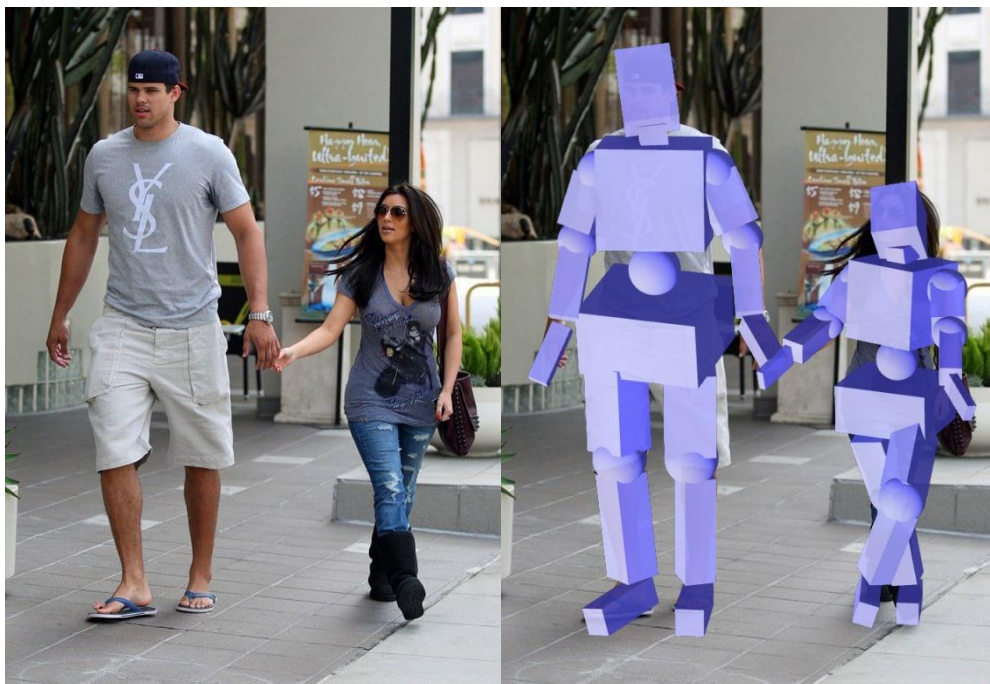


Рисунок 2.27 – Результат работы программы для 2-ух людей

И в-третьих, алгоритм успешно подстраивается к телосложению людей (рисунки 2.28-2.29).



Рисунок 2.28 – Результат работы программы



Рисунок 2.29 – Результат работы программы

Таким образом, программа получает на вход фотографию человека. Далее накладывает на исходное изображение построенную модель и сохраняет в формате «.png», а также сохраняет трехмерный объект в формате «.obj». Вторая функция позволяет при желании изменить ракурс обзора на позу человека с фото (рисунок 2.30).

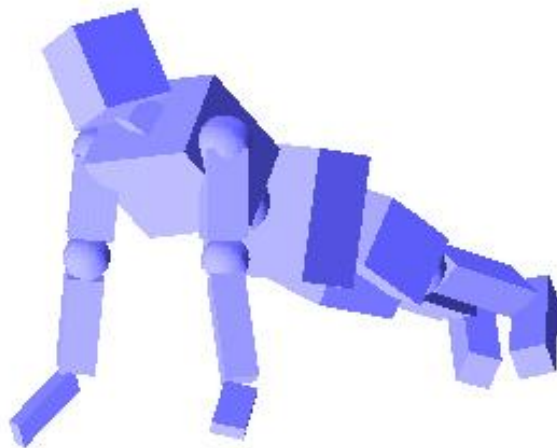


Рисунок 2.30 – Сохранение объекта в формате «.obj»

Кроме всего вышеизложенного, данная методика позволяет находить ошибки на уже готовых рисунках, когда неопытному человеку сложно их увидеть. Так, например, рассмотрим работу на рисунке 2.31. Как можно заметить, рисунок получился плоским и

статичным. Но наложенная вспомогательная конструкция помогает обнаружить недочеты, а именно: грудная клетка на исходном фото несколько наклонена назад, а голова смотрит вверх, чего не хватает в иллюстрации.



Рисунок 2.31 – Сравнение рисунка по фото и результата работы программы

Можно сделать вывод, что данное программное обеспечение можно использовать в художественных школах и при самообучении для развития пространственного мышления и тренировки художественных навыков.

ЗАКЛЮЧЕНИЕ

В выпускной квалификационной работе автором была рассмотрена узкоспециализированная задача визуализации человеческой фигуры в виде простых геометрических тел, а именно: параллелепипедов. Данная модель является вспомогательным инструментом для художников, упрощающим восприятие формы и конструкции изображаемого объекта.

1) Для достижения поставленной цели определены требования к модели, основанные на эмпирическом анализе предметной области.

2) Изучены различные модели, методы и алгоритмы восстановления трехмерных объектов по фотографии. На основе анализов производительности и удовлетворения требуемым критериям, выбрана подходящая модель оценки позы по фотоснимку, а также определены входные данные для дальнейшего развития алгоритма.

3) Сформулировано математическое описание искомой конструкции и разработано решение задачи.

4) Также предусмотрены возможные возмущения на входе в виде выступов на поверхности исходного объекта и выбрана концепция, нечувствительная к таким неровностям. Для этого введена аппроксимация шестиугольными фигурами, выбранная в связи с тем фактом, что природные тела, а именно человеческое тело, имеет скругленные контуры, далекие от идеальных параллелепипедов.

5) Оценено оптимальное количество окаймляющих трехмерных тел для создания модельной фигуры приемлемого восприятия. Данный алгоритм назван «Задачей об окаймляющем параллелепипеде», который может быть использован как в двумерном, так и в трехмерном пространствах.

6) Также введена эвристика поворота множества, требующая на входе приблизительно симметричного расположения точек относительно диагоналей. Угол поворота рассчитывается на основании разницы сумм в разных координатных четвертях и в разных областях относительно границы фигуры шестиугольника.

7) Разработано ПО и проведено тестирование.

8) Основные результаты проделанной работы опубликованы в тезисах студенческой конференции «75-е Дни науки студентов МИСиС» [35].

Алгоритм показал достойные результаты и воспроизвел изображения, очень близкие к материалам из учебных пособий для художников.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1 Золотарева Л. Р. "Художественная картина мира" итальянского Ренессанса в контексте освоения истории искусства // Известия АлтГУ. Филология и искусствоведение. 2013. – С. 182-187.

2 Бакиева О. А. Методика преподавания изобразительного искусства: Учебное пособие. – Тюмень: Издательство Тюменского государственного университета, 2012.

3 Запорожец А.В. Сенсорное воспитание дошкольников / А. В. Запорожец, А.П. Усова. – М.: Академия, 1963.

4 Каплунович И. Я. Формирование структуры пространственного мышления учащихся при решении математических задач: Дисс. ... канд. пед. наук. - М., 1978.

5 Власов В. Г. Обращаемое пространство в изобразительном искусстве и проективном мышлении [Электронный ресурс] // Архитектон: известия вузов. 2019. №1(65). URL: http://old.archvuz.ru/PDF/%23%2065%20PDF/%2365_20_Vlasov.pdf (дата обращения: 25.02.2020).

6 Белякова Л. Г. Теоретические основы обучения пространственному рисованию // Вестник Вятского государственного гуманитарного университета. Педагогика и психология. – 2010.– № 2(3). – С. 50-55.

7 Молева Н. М., Белютин Э. М. Школа Антона Ашбе. К вопросу о путях развития художественной педагогики на рубеже XIX-XX веков. – М.: Искусство, 1958.

8 Черниенко И. Б. Научное мировоззрение художника эпохи Возрождения (на примере Альбрехта Дюрера). // Вестник Пермского университета. История. – 2004. – № 5. – С. 51-54.

9 Павлова М. А. Метод «поверочного рисования» П. П. Чистякова в современном художественном образовании // Евразийский Союз Ученых (ЕСУ). Педагогические науки. – 2014. – № 9. – С. 46.

10 Сапожников А. П. Полный курс рисования: Учеб. пособие. Под редакцией Ларионова В. Н. – М.: АЛЕВ-В, 2003.

11 Мак-Каллок У. С., Питтс В. Логическое исчисление идей, относящихся к нервной активности // Автоматы / Под ред. К. Э. Шеннона и Дж. Маккарти. – М.: Изд-во иностр. лит., 1956.

12 Зенин А. В. Исследование возможностей использования нейронных сетей. – Казань: Молодой ученый, 2017.

13 Гудфеллоу Я., Глубокое обучение.: Учеб. пособие. // Гудфеллоу Я., Бенджио И., Курвилль А. – М. Изд-во ДМК-Пресс, 2018.

- 14 Форсайт Д. А., Понс Ж. Компьютерное зрение. Современный подход. – М.: Вильямс, 2004.
- 15 Васильков Д. М. Геометрическое моделирование и компьютерная графика вычислительные и алгоритмические основы: курс лекций. – Минск: БГУ, 2011.
- 16 Chan T. M. Output-Sensitive Construction of Convex Hulls: Ph.D. thesis / Department of Computer Science. – University of British Columbia, 1995.
- 17 Jarvis A. On the identification of the convex hull of a finite set of points in the plane. // In-formation Processing Letters. – 1973.
- 18 Andrew A. M. Another efficient algorithm for convex hulls in two dimensions // Information Processing Letters. – 1979.
- 19 Graham R. L. An efficient algorithm for determining the convex hull of a finite planar set // Information Processing Letters. – 1972.
- 20 Huynh J. Separating Axis Theorem for Oriented Bounding Boxes [Electronic resource] // jkh.me: 2015.
URL: jkh.me/files/tutorials/Separating%20Axis%20Theorem%20for%20oriented%20Bounding%20Boxes.pdf (дата обращения: 14.10.2019).
- 21 Ericson C. Real-Time Collision Detection // Morgan Kaufmann. – 2004.
- 22 Файзрахманов Р. А. Алгоритмы обнаружения коллизий плоских двумерных объектов произвольной формы 2015 // Файзрахманов Р. А., Мурзакаев Р. Т., Швецов М. Д., Мехоношин А. С., Хабибулин А. Ф. [Электронный ресурс]. – Электронный научный журнал «Инженерный вестник Дона», 2015
- 23 Лесков А. Г., Селиверстова Е. В. Расчет областей пересечения поверхностей захватных устройств манипуляторов и деформируемых объектов при планировании и моделировании захвата. // Вестник МГТУ им. Н. Э. Баумана. Сер. Приборостроение. – 2016.
- 24 Роджерс Д., Адамс Дж. Математические основы машинной графики. – М.: Мир, 2001.
- 25 Chowdhary C. L. Appearance-based 3-D Object Recognition and Pose Estimation: Using PCA, ICA and SVD-PCA Techniques. – LAP LAMBERT Academic Publishing. – 2011.
- 26 Alexander Toshev, Christian Szegedy, DeepPose: Human Pose Estimation via Deep Neural Networks. — CA, US. – 2014.
- 27 Bishop C. M. Pattern Recognition and Machine Learning // Journal of electronic imaging. – 2006.
- 28 Коэльо Л., Ричарт В. Построение систем машинного обучения на языке Python – М.: ДМК Пресс. – 2016.

29 Bogo F. Keep it SMPL: automatic estimation of 3d human pose and shape from a single image. // Bogo F., Kanazawa A., Lassner C., Gehler P. V., Romero J., and Black M. J. – In: ECCV. –2016.

30 Pishchulin, L. DeepCut: Joint subset partition and labeling for multi person poseestimation. // Pishchulin, L., Insafutdinov, E., Tang, S., Andres, B., Andriluka, M., Gehler, P., Schiele, B. – In: IEEE Conference on Computer Vision and Pattern Recognition, CVPR. – 2016.

31 Loper M. A skinned multi-person linear model. // Loper M., Mahmood N., Romero J., Pons-Moll G., and Black M. J. – ACM Trans. Graphics (Proc. SIGGRAPH Asia). –2015.

32 Anguelov D. SCAPE: shape completion and animation of people. // Anguelov D., Srinivasan P., Koller D., Thrun S., Rodgers J., and Davis J. – ACM transactions on graphics (TOG). – 2005

33 Botsch M. Geometric Modeling Based on Triangle Meshes. // Botsch M., Pauly M., Kobbelt L., Pierre Alliez P., Lévy B. – ACM (SIGGRAPH). – 2006.

34 Букатов А. А., Гридчина Е. Е., Заставной Д. А. Методы скелетной анимации для трансформации полигональных поверхностей трёхмерных моделей. // Электронный научный журнал «Инженерный вестник Дона». 2012. URL: http://www.ivdon.ru/uploads/article/pdf/2012_3_11.pdf_897.pdf (дата обращения: 02.04.2019).

35 Цой Д. А. Алгоритм декомпозиции человеческой позы на элементы простой формы по фотоснимку // 75-е дни науки студентов НИТУ «МИСиС» международные, межвузовские и институтские научно-технические конференции Книга 2. – М.: Издательский Дом МИСиС, 2020. - С. 827.