

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г.ЧЕРНЫШЕВСКОГО»

Кафедра системного анализа и
автоматического управления

**ОЦЕНКА СХОЖЕСТИ ТРЕХМЕРНЫХ ИЗОБРАЖЕНИЙ НА ОСНОВЕ
АНАЛИЗА ИХ ПОВЕРХНОСТЕЙ**

АВТОРЕФЕРАТ МАГИСТЕРСКОЙ РАБОТЫ

студента 2 курса 273 группы
направления 02.04.03 Математическое обеспечение и администрирование
информационных систем
факультета компьютерных наук и информационных технологий
Трофимука Павла Викторовича

Научный руководитель

доцент, к.ф.-м.н.

М.В. Корнилов

Зав. кафедрой

к.ф.-м.н., доцент

И.Е. Тананко

Саратов 2025

ВВЕДЕНИЕ

Актуальность темы. Сравнение трехмерных изображений сопряжено с рядом трудностей. В отличие от двумерных изображений, 3D-изображения содержат сложную пространственную информацию, включая геометрию, топологию и текстуру. Представление поверхностей включает в себя сложные структуры данных и математические модели, которые отражают нюансы форм в трех измерениях. Кроме того, изображения могут различаться по масштабу, ориентации и расположению, что требует надежных методов сравнения, инвариантных к таким преобразованиям.

Дополнительная сложность связана с разнообразием форматов, используемых для хранения графических данных. Различные форматы файлов, каждый со своими спецификациями и ограничениями, могут по-разному представлять 3D-поверхности. Такая неоднородность усложняет процесс стандартизации методов сравнения и требует всестороннего понимания этих форматов для обеспечения точных оценок.

Несмотря на то, что для сравнения трехмерных изображений существует множество сложных алгоритмов, существует потребность в более простых методах, которые сочетают эффективность с приемлемой точностью. Такие методы особенно ценны в приложениях с ограниченными вычислительными ресурсами, где достаточно аппроксимации.

Целью магистерской работы является изучение и программная реализация оценки схожести трехмерных изображений на основе анализа их поверхностей.

Поставленная цель определила **следующие задачи:**

1. Подбор и изучение литературы по принципам построения поверхностей трехмерных изображений, по основным форматам хранения графических данных, по простейшим методам сравнения изображений.

2. Программная реализация алгоритма представления поверхности трехмерной модели в виде регулярной сетки, а также триангуляции Делоне. Сравнение эффективности работы алгоритмов

3. Программная реализация алгоритма сравнения изображений. Подготовка и проведение вычислительных экспериментов.

4. Сбор, обработка и анализ результатов, полученных в ходе экспериментов для оценки схожести изображений на основе анализа их поверхностей при различных наборах данных.

Методологические основы сравнения трехмерных изображений на основе анализа их поверхностей представлены в работах Н.Ф. Дышкант [1], А.В. Меженина [2], J.F. Hughes [3], В. Cyganek [4], L. Luciano [5], G. Yona [6].

Теоретическая значимость магистерской работы в систематизации представлений о поверхностях 3D-объектов: от параметрических (Безье, В-сплайн, NURBS) и неявных моделей до полигональных сеток и алгоритмов их реконструкции из облаков точек.

Практическую значимость магистерской работы представляет стек алгоритмов — от базовых (регулярная сетка, триангуляция Делоне) до продвинутых (ICP, HKS/WKS/AISIHK, PointNet) и их гибридного сочетания. Применяя гибридный алгоритм в сравнительном эксперименте на открытых наборах ModelNet40 и FAUST, демонстрируются преимущества подобного подхода и потенциал для дальнейшего изучения, улучшения и применения результатов в отдельных областях.

Структура и объем работы. Магистерская работа состоит из введения, 5 разделов, заключения, списка использованных источников и 8 приложений. Общий объем работы – 115 страниц, из них 65 страниц – основное содержание, включая 7 рисунков и 3 таблицы, список использованных источников информации – 22 наименования.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Первый раздел «Принципы построения поверхностей трехмерных изображений» посвящен формированию теоретического фундамента по задаче сравнения трехмерных моделей. Глава начинается с определения поверхности как двумерного многообразия в R^3 и показывает два базовых способа её задания. Параметрическая поверхность описывается отображением

$$S(u, v) = (x(u, v), y(u, v), z(u, v)), \quad (u, v) \in \Omega \subset R^2, \quad (1.1)$$

что обеспечивает локальный контроль формы и удобство при редактировании, тогда как неявная поверхность задаётся уравнением

$$F(x, y, z) = 0, \quad (1.2)$$

позволяя естественно работать с меняющейся топологией.

Основной объём раздела посвящён математическим представлениям. Через формулы

$$C(t) = \sum_{i=0}^n B_i^n(t) P_i, \quad (1.3)$$

$$S(u, v) = \sum_{i=0}^n \sum_{j=0}^m B_i^n(u) B_j^m(v) P_{i,j}, \quad (1.4)$$

раскрыты кривые и поверхности Безье, после чего вводятся В-сплайны:

$$C(t) = \sum_{i=0}^n N_{i,k}(t) P_i, \quad (1.5)$$

и рациональные NURBS

$$C(t) = \frac{\sum_{i=0}^n N_{i,k}(t) w_i P_i}{\sum_{i=0}^n N_{i,k}(t) w_i}, \quad (1.6)$$

обеспечивающие точное моделирование и локальное редактирование формы за счет весов.

Переходя к дискретизации, рассматриваются полигональные сетки – стандарт обмена 3D-содержимым благодаря аппаратной поддержке графических процессоров. Подробно анализируются структуры данных, критерии качества и алгоритмы оптимизации. Особое внимание уделяется триангуляции Делоне, которая максимизирует минимальный угол треугольников и служит основой для конечно-элементных расчётов и оценки сходства моделей.

Раздел завершается обзором реконструкции поверхности из облаков точек, получаемых лазерным сканированием и фотограмметрией. Описаны этапы фильтрации, сглаживания и алгоритм Пуассоновской реконструкции.

Выводы раздела:

1. Выбор представления поверхности определяет класс допустимых метрик и алгоритмов сравнения.

2. Достоверная реконструкция из облака точек является необходимым условием корректной оценки сходства трёхмерных изображений.

Второй раздел «Основные форматы хранения графических данных» посвящен популярным форматам и тому, как формат файла влияет на точность, объём и скорость дальнейшей обработки.

Сначала рассматриваются возможности форматов по хранению геометрии (вершины, ребра и грани), дополнительных атрибутов (нормали, UV-координаты, материалы), анимации и метаданных. Отдельно анализируется компромисс между размером и качеством (сжатие без/с потерями), поддержка потоковой загрузки и кроссплатформенность.

Далее идёт подробный обзор наиболее частых форматов. Текстовый OBJ удобен для отладки, содержит списки вершин, текстур, нормалей и полигонов, но не поддерживает иерархии и анимацию, а при большом числе полигонов быстро разрастается в мегабайты. STL, напротив, описывает только треугольные грани и благодаря простоте стал де-факто стандартом 3D-печати, однако отсутствие материалов ограничивает его в игровых и VR-проектах.

Форматы COLLADA и FBX предназначены для насыщенных сцен: оба умеют хранить скелетные анимации, уровни детализации и физические свойства; COLLADA открыт и основан на XML, FBX – двоичный, компактный и поддерживается большинством коммерческих движков.

Формат хранения задаёт «верхнюю границу» точности и производительности всего последующего конвейера: он определяет, какие атрибуты формы сохранятся, как быстро модель загрузится и сможет ли алгоритм получить доступ к высокоуровневой семантике.

Третий раздел «Простейшие методы сравнения изображений» переходит от описания данных к базовым метрикам геометрического сходства и содержит две собственные реализации.

Для дискретных облаков точек рассматриваются две классические

дистанции: чувствительное к выбросам расстояние Хаусдорфа и более устойчивое Chamfer

$$C(A, B) = \frac{1}{|A|} \sum_{a \in A} \min_{b \in B} |a - b| + \frac{1}{|B|} \sum_{b \in B} \min_{a \in A} |b - a|, \quad (1.7)$$

которое усредняет ошибки и лучше масштабируется на больших выборках.

Для алгоритма регулярной сетки написан скрипт, превращающий модель в воксельное облако и оценивающий пересечение над объединением

$$\text{IoU} = \frac{\text{vol}(V_1 \cup V_2)}{\text{vol}(V_1 \cap V_2)}, \quad (1.8)$$

что даёт быструю и устойчивую к поворотам оценку «скелетного» сходства форм. Код нормализует модели, дискретизирует их с выбранным размером вокселя и замеряет время, память и загрузку CPU во время сравнения. Результаты показали, что метод пригоден для предварительной фильтрации кандидатов и обработки больших коллекций.

Вторая разработанная программа строит тетраэдрическое разбиение объёма каждой модели, извлекает данные (распределения длин рёбер, углов, объёмов тетраэдров) и сравнивает их по расстоянию Бхаттачарьи и пересечению гистограмм. Несмотря на большие затраты ресурсов, подход выявляет более тонкие различия формы, недоступные вокселизации, и демонстрирует возросшую возможность различать модели.

Базовые метрики уже позволяют разделять модели по крупным геометрическим признакам, но чувствительны к шуму и не учитывают семантическое сходство. В то же время реализованные прототипы показали, что грамотная оптимизация (downsampling, KD-деревья, параллельные вычисления) делает даже простые методы пригодными для оперативной предобработки больших наборов данных.

Четвертый раздел «Методы и алгоритмы сравнения трехмерных изображений» начинается с постановки задачи: после грубого геометрического сравнения требуется метод, чувствительный к пластике формы и одновременно стойкий к деформациям. Поэтому были отобраны и реализованы три алгоритма.

Первым изложен классический Iterative Closest Point (ICP). Алгоритм минимизирует функцию

$$\min_{R,t} \sum_{i=1}^N |Rp_i + t - q_i|^2, \quad (1.9)$$

подбирая вращение R и сдвиг t между облаками $\{p_i\}$ и $\{q_i\}$. В реализации модели предварительно вокселизуются с шагом $v = 0.01$, нормали оцениваются через гибридный поиск по KD-дереву, используя RANSAC для ускорения сходимости.

Далее рассматриваются спектральные дескрипторы HKS/WKS/AISIHK. Основой служит тепловое ядро

$$k_t(x, y) = \sum_j e^{-\lambda_j t} \phi_j(x) \phi_j(y), \quad (1.10)$$

чья диагональ образует дескриптор

$$\text{HKS}_t(x) = k_t(x, x) = \sum_j e^{-\lambda_j t} \phi_j(x)^2, \quad (1.11)$$

Скрипт строит k -NN-граф ($k=20$), выбирает 128 логарифмически равных «вреён», что делает дескриптор одновременно локальным и глобальным. Дополнительно введена нормировка времени, устраняющая зависимость от масштабирования модели.

Третьим рассматривается PointNet — нейросетевой эмбеддер облаков точек. После аугментирования (шум, случайное вращение, масштаб) данных, сеть выводит 1024-мерный признак z , который нормируется

$$\hat{z} = \frac{z}{|z|_2}, \quad (1.12)$$

а косинусная близость

$$\text{sim}(i, j) = \hat{z}_i \cdot \hat{z}_j, \quad d = 1 - \text{sim}, \quad (1.13)$$

служит дистанцией.

Последним в разделе описывается гибридный алгоритм — позднее слияние трёх независимых оценок:

$$S = \alpha S_{\text{ICP}} + \beta S_{\text{HKS}} + \gamma S_{\text{PN}}, \quad \gamma = 1 - \alpha - \beta, \quad (1.14)$$

Параметры по умолчанию $\alpha=0.3$, $\beta=0.2$ подобраны перекрёстной

проверкой; показано, что рост α улучшает точность на жёстких CAD-моделях, тогда как увеличение β выгодно для гибких сканов человеческих поз.

Комплекс из ICP, спектральных дескрипторов, PointNet-эмбединга и их взвешенной комбинации покрывает жесткую геометрию, не жесткие преобразования и преимущества семантического подхода.

Пятый раздел «Сравнение эффективности алгоритмов» посвящён эмпирической верификации методов на двух эталонных наборах: ModelNet40 [7] (классические CAD-объекты) и FAUST [8] (сканы людей в разных позах). Чтобы соотнести машинные оценки с человеческим восприятием, используются три группы метрик. Для задачи поиска формулируются

$$P_q(k) = \frac{1}{k} \sum_{i=1}^k 1[l_{r_i} = l_q], \quad AP(q) = \frac{1}{R_q} \sum_k P_q(k) \Delta R_q(k), \quad (1.15)$$

а затем усредняются в

$$mAP = \frac{1}{N} \sum_q AP(q), \quad (1.16)$$

Качество ранжирования также оценивается ROC-AUC

$$AUC = \Pr(S_{ij} > S_{mn} \mid y_{ij} = 1, y_{mn} = 0), \quad (1.17)$$

Наконец, структурность всей матрицы сходства проверяется внешними (NMI, ARI) и внутренними (силуэт, Дэвис-Боулдин) метриками, связанными с кластеризацией.

На ModelNet40 гибридный метод показал $mAP=0.410$, $ROC-AUC = 0.800$ и $P@1=0.464$, опередив ICP и PointNet по каждому критерию. При этом ICP страдает от несоответствия поз и даёт худший NMI, тогда как PointNet недооценивает мелкие детали, снижая $P@10$ до 0.146.

На FAUST картина иная: спектральный HKS уже сам по себе достигает $P@1=0.501$ и $mAP=0.432$ благодаря изометрической инвариантности. Тем не менее гибрид, добавив PointNet-контекст, поднимает mAP до 0.453 и удерживает $ROC-AUC = 0.806$, что эквивалентно правильной сортировке в 80% случаев.

Комплексные показатели подтверждают вывод: гибрид минимизирует

ложные срабатывания — NMI растёт до 0.83, силуэт-коэффициент до ≈ 0.25 , а индекс Дэвиса-Боулдина падает ниже 1.30, что указывает на явные и компактные кластеры.

Наборы экспериментов демонстрируют, что ни одна одиночная стратегия не универсальна: ICP хорош для твёрдых объектов, HKS — для объектов, сохраняющих общую структуру, PointNet — для семантики. Разработанный гибрид в большинстве случаев превосходит алгоритмы в отдельности, увеличивая mAP на $\approx 4\text{--}6\%$ и стабилизируя все качественные метрики.

ЗАКЛЮЧЕНИЕ

Исследование трехмерных методов визуализации и моделирования выявило обширный ландшафт методологий, проблем и достижений, которые играют решающую роль в областях компьютерной графики, компьютерного зрения и распознавания образов. В данной работе были подробно рассмотрены принципы построения поверхностей 3D-изображений, основные форматы хранения графических данных, простейшие методы сравнения изображений. Был реализован набор алгоритмов и решена задача сравнения трёхмерных моделей по их поверхностям на разных наборах данных (выборке CAD-объектов и сканов с реалистичными искажениями: шум, неполное покрытие, изометрические деформации).

Предложенный гибрид демонстрирует устойчивость к небольшому шуму, частичным сканам и деформациям, обеспечивая стабильно высокую точность поиска похожих моделей на датасетах ModelNet и FAUST в сравнении с каждым из базовых алгоритмов (ICP, PointNet, HKS) по отдельности.

Основные источники информации:

- 1 Дышкант, Н. Ф. Эффективные алгоритмы сравнения поверхностей, заданных облаками точек : дис. ... канд. физ.-мат. наук : 01.04.03 / Н. Ф. Дышкант. — М., 2011. — 139 с.
- 2 Меженин, А. В. Технологии разработки 3D-моделей : учеб. пособие / А. В. Меженин. — СПб : Университет ИТМО, 2018. — 100 с.

- 3 Hughes, J.F. Computer Graphics : Principles and Practice / J.F. Hughes [и др]. — 3rd ed. — Boston : Addison-Wesley, 2014. — 1264 p.
- 4 Cyganek, B. An Introduction to 3D Computer Vision Techniques and Algorithms / B. Cyganek, J. P. Siebert. — Chichester : Wiley, 2009. — 483 p.
- 5 Luciano, L. Deep Similarity Network Fusion for 3D Shape Classification / L. Luciano, A. Ben Hamza // *The Visual Computer*. — 2019. — Vol. 35, № 5. — P. 1171-1180.
- 6 Yona, G. Neural Descriptors: Self-Supervised Learning of Robust Local Surface Descriptors Using Polynomial Patches [Электронный ресурс] / G. Yona [и др]. // arXiv [Электронный ресурс] : электронный архив — 2025. — URL: <https://arxiv.org/abs/2503.03907> (дата обращения: 10.03.2025). — Загл. с экрана. — Яз. англ.
- 7 Wu, Z. 3D ShapeNets: A Deep Representation for Volumetric Shapes / Z. Wu [и др]. // Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). — Boston, 2015. — P. 1912-1920.
- 8 Bogo, F. FAUST: Dataset and Evaluation for 3D Mesh Registration / F. Bogo [и др.]. // Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). — Columbus, 2014. — P. 3794-3801.