

# СОЗДАНИЕ ЦИФРОВЫХ ДВОЙНИКОВ СОЧЛЕНЕННЫХ ОБЪЕКТОВ НА ОСНОВЕ ИХ ВЗАИМОДЕЙСТВИЯ

Тимаков Кирилл Александрович

МИРЭА — Российский технологический университет

Tim1997.10@yandex.ru

## CREATING DIGITAL DOUBLES OF ARTICULATED OBJECTS BASED ON THEIR INTERACTION

*K. Timakov*

*Summary.* Digitization of physical objects in the digital space has the potential to open new research and applications in embodied AI and mixed reality. This work examines the issue of creating interactive digital counterparts of real articulated objects that can be used in a virtual environment. The article also examines the Ditto model for studying the evaluation of articulated objects and reconstructing the three-dimensional geometry of an articulated object through interactive perception based on the data obtained. Given a pair of visual observations of the object of interest before and after the interaction, the Ditto model reconstructs the geometry at the level of the parts and evaluates the articulation model of the object. Based on this model, the use of the Iterative Closest Point (ICP) algorithm is proposed, which significantly improves the quality of the input data to improve the quality of the resulting digital doubles and correctly determine their articulation.

*Keywords:* Ditto, ICP, point clouds, hinge, articulated object, AI, digital twin.

*Аннотация.* Оцифровка физических объектов в цифровом пространстве имеет потенциал для открытия новых исследований и приложений в воплощенном ИИ и смешанной реальности. Данная работа рассматривает вопрос создания интерактивных цифровых двойников реальных сочлененных объектов, которые могут быть использованы в виртуальной среде. Так же в статье идет рассмотрение модели Ditto для изучения оценки шарнирных объектов и реконструкции трехмерной геометрии сочлененного объекта посредством интерактивного восприятия на основе полученных данных. Учитывая пару визуальных наблюдений объекта интереса до и после взаимодействия, модель Ditto реконструирует геометрию на уровне частей и оценивает модель сочленения объекта. На основе данной модели предлагается использование алгоритма Iterative Closest Point (ICP), что позволяет значительно улучшить качество входных данных, для улучшения качества получаемых цифровых двойников и корректного определения их сочленения.

*Ключевые слова:* Ditto, ICP, облака точек, шарнир, сочлененный объект, ИИ, цифровой двойник.

## Введение

Синтетические данные приобретают все большее значение в развитии современных приложений искусственного интеллекта, обеспечивая необходимые ресурсы для обучения и прототипирования моделей компьютерного зрения, а также для обучения роботизированных систем выполнению физических задач. В условиях экспоненциального роста масштаба и требовательности к данным современных моделей ИИ, виртуальные платформы и синтетические наборы данных предоставляют экономически эффективный и масштабируемый источник обучающих материалов. Для обеспечения эффективного использования синтетических данных в обучении моделей машинного зрения, критически важным фактором является реализм, поскольку расхождения в распределениях данных между реальным и виртуальным мирами ограничивают обобщающую способность моделей, обученных в симуляционных средах. Одним из перспективных направлений для уменьшения разрыва в реализме является оцифровка физических объектов и их последующее воссоздание в цифровых средах. Несмотря на значительные

достижения в области 3D-зрения и одновременной локализации, и картографирования (SLAM) в части захвата реалистичных объектов и сцен с формированием статических трехмерных моделей, растущий спрос со стороны исследований в области воплощенного искусственного интеллекта и смешанной реальности требует разработки интерактивных цифровых двойников физических объектов, способных к интеграции в смоделированные среды и взаимодействию с виртуальными агентами. В частности, создание цифровых двойников сочлененных объектов представляет собой сложную научно-техническую задачу. Процесс работы визуализации 3D с использованием ИИ модели Ditto представлен на рисунке 1.

Рассматриваемый метод создания цифровых двойников сочлененных объектов, основанный на интерактивном восприятии, позволяет совместно реконструировать геометрию отдельных элементов и оценить модель сочленения объекта, используя визуальные данные, полученные до и после взаимодействия. Полученные в результате цифровые двойники могут быть интегрированы в физические движки и обеспечивают полную интерактив-

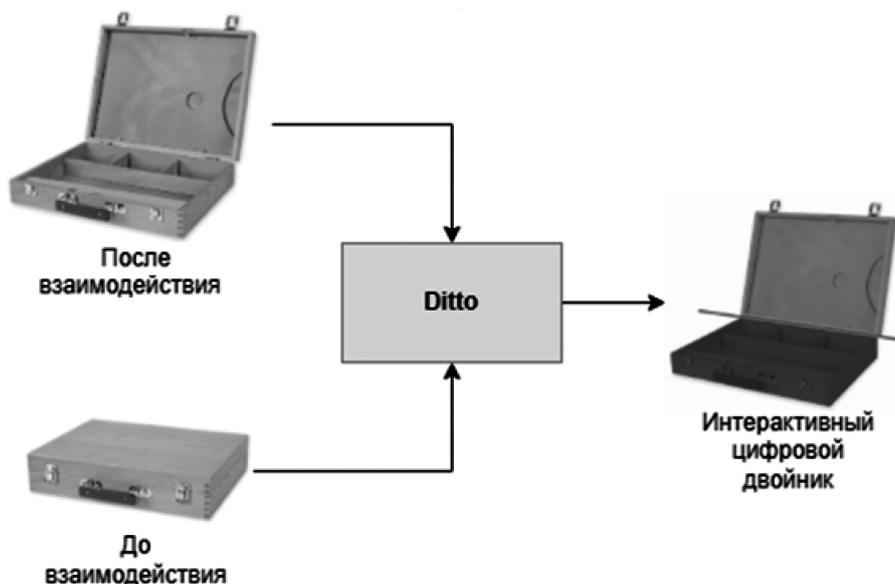


Рис. 1. Представление работы визуализации 3D с использованием ИИ модели Ditto

ность в приложениях моделирования движения роботов и/или дополненной/виртуальной реальности (AR/VR).

Рассматривая ситуацию цифровых двойников, можно заметить, что в последние годы в рамках развития воплощенных платформ искусственного интеллекта наблюдается тенденция интеграции интерактивных сочлененных или шарнирных объектов («articulated objects»), к примеру, предметы быта, в симулированные с целью обучения виртуальных агентов. Существующие методы создания в значительной степени зависят от участия ряда специалистов, что ограничивает масштабируемость систем. Автоматизация оценки и реконструкции сочлененных объектов на основе методов машинного зрения является активно развивающимся направлением исследований, которому способствуют новые инструменты, разработанные сообществом 3D-зрения, в частности методы геометрического глубокого обучения и неявные нейронные представления.

В рамках данной работы поставлена задача рассмотрение полноценных виртуальных моделей сочлененных физических объектов, не принадлежащих к известным категориям, а также рассмотрения алгоритма их визуализации для дальнейшего улучшения с помощью применения нового подхода с использованием алгоритма ICP. Данные цифровые двойники должны представлять геометрические и физические характеристики отдельных частей объекта, а также параметры их сочленений, включая призматические и вращательные шарниры. Оценка параметров сочленения по одному изображению, независимо от категории объекта, является неоднозначной задачей, поскольку движение частей может интерпретироваться как перемещение вдоль призматической оси или вращение вокруг оси вращения, в зависимости от кинематических характеристик

шарнира. В продолжение исследований в области интерактивного восприятия шарнирных объектов, предлагается метод построения цифровых двойников на основе визуальных данных, полученных до и после осуществления движения сочлененных частей. Данная модель включает в себя три взаимосвязанные подзадачи: сегментацию объекта на отдельные части с использованием информации о движении, реконструкцию геометрии частей на основе неполных данных в виде облака точек, и оценку параметров сочленения с учетом неизвестного типа сочленения [1]. На рисунке 2, представлена схема взаимодействия между различными этапами решения задачи создания цифровых двойников сочлененных объектов.

В данной работе рассматривается система Ditto (Digital twin of articulated objects), представляющая собой неявную нейронную модель, разработанную для совместного предсказания геометрии отдельных частей и кинематических параметров сочленений. Для кодирования трехмерной информации с высокой разрешающей способностью и непрерывностью используются неявные нейронные представления. Архитектура Ditto основана на ConvONets, которая обучает локальные неявные поля, используя сверточные сетки признаков. Входными данными для Ditto являются частичные данные в виде облаков точек, представляющих сочлененный объект до и после взаимодействия с одной из его частей. Ключевой технической задачей является установление соответствия между этими двумя наборами данных. Для этого облака точек кодируются с использованием PointNet++ в два набора подвыборочных точечных признаков. Далее, эти наборы объединяются посредством слоя самовнимания, после чего объединенные подвыборочные признаки декодируются в плотные точечные признаки. На основе декодированных

Задачи создания цифровых двойников articulated objects

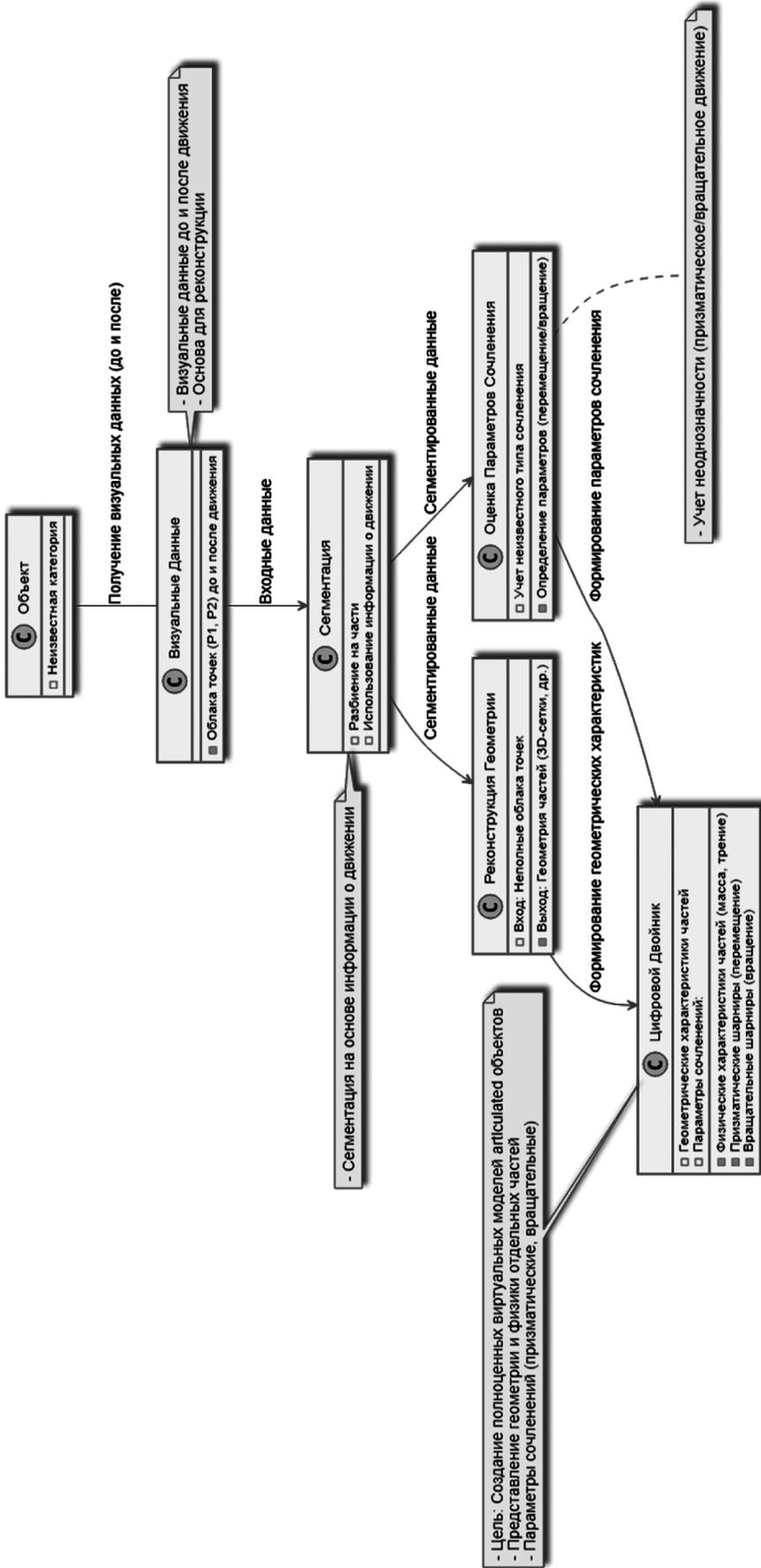


Рис. 2. Схема последовательности действий и взаимосвязи между различными этапами решения задачи создания цифровых двойников сочлененных объектов

точечных признаков строится структурированная сетка признаков, что позволяет вычислять локальный признак для заданной трехмерной координаты запроса. Система включает неявный декодер занятости и декодер сегментации, которые отображают трехмерную координату и ее локальный признак в метку занятости/сегментации части, обеспечивая воссоздание геометрии на уровне отдельных частей. Также, использование набора неявных декодеров, предсказывают относительные параметры сочленения в каждой точке запроса. Такой подход к предсказанию параметров сочленения позволяет повысить надежность оценки по сравнению с глобальным предсказанием параметров сочленения.

### Формулировка проблемы

В данной работе изучается и рассматривается решение проблемы воссоздания интерактивных цифровых двойников шарнирных объектов из пары изображений объекта до и после взаимодействия, а также способы улучшения создания таких двойников при помощи различных алгоритмов и моделей. Цифровые двойники обычно представлены в стандартных 3D-форматах, таких как URDF, чтобы их можно было импортировать в физические движки. Чтобы обеспечить физическое взаимодействие в виртуальном мире, цифровой двойник шарнирного объекта представляет собой кинематическое дерево, где узлы определяют геометрию и физические свойства (например, массу и трение) отдельных частей, а ребра определяют кинематические соединения между частями. Эта работа фокусируется на оценке геометрии частей объекта и кинематики шарниров при установке физических свойств на значения по умолчанию на основе собранной информации. На рисунке 3, представлена архитектура ИИ модели Ditto, которая предлагается для использования с целью зарождения мест сочленения объекта.

После этапа кодирования с использованием PointNet++ [2], производится слияние подвыборочных точечных признаков с помощью простого слоя внимания. Далее, используются два независимых декодера для преобразования объединенных точечных признаков в два набора плотных точечных признаков, предназначенных для реконструкции геометрии и оценки параметров сочленения, соответственно. Посредством проецирования и усреднения точечных признаков конструируются сетки/плоскости признаков, что позволяет извлекать локальные признаки для заданных координат запроса. С учетом локальных признаков различные декодеры используются для предсказания занятости, сегментации и параметров шарниров относительно точек запроса.

Моделью предполагается, что у объекта есть только одна подвижная часть, с которой возможно взаимо-

действие. Входными данными для Ditto являются пара наблюдаемых облака точек  $P1, P2 \in \mathbb{R}^{N \times 3}$  сочлененного объекта до и после взаимодействия, где  $N$  — количество входных точек. Конечной целью работы модели является сегментация и реконструкция трехмерной геометрии для статических подвижных частей, оценка параметров соединений, двух частей, и относительное изменение состояний соединений. Стоит отметить, что представленные методы могут работать и для обычных не подвижных частей для их 3D визуализации, что даст в дальнейшем возможность подумать над реализацией подвижных частей в данных объектах.

Для оценки параметров сочленения рассматриваются одномерные вращательные (revolute) и призматический (prismatic) шарниры. Используется параметризация шарниров, предложенная в работе [3]. Параметры поступательного шарнира включают в себя направление оси перемещения  $u_p \in \mathbb{R}^3$  и состояние шарнира  $s^p$ . Состояние шарнира  $s^p$  определяется как относительное расстояние перемещения между двумя наблюдениями. Параметры вращательного шарнира включают в себя направление оси вращения  $u^r \in \mathbb{R}^3$ , опорную точку  $q \in \mathbb{R}^3$  на оси вращения и состояние шарнира  $s^r$ . Состояние шарнира  $s^r$  определяется как относительный угол поворота между двумя наблюдениями.

### Двухпоточный кодер

Для реализации совместного обучения трехмерной реконструкции и оценки модели сочленения используется извлечение признаков, объединяющих информацию из входной пары облаков точек. Для построения кодировщика используется архитектура ConvONets, представляющая собой современный метод 3D реконструкции, основанный на неявном представлении [4]. Для объединения двух наборов точечных признаков, полученных из двух наборов облаков точек, применяется слой внимания. Следует отметить, что вычислительная сложность операций слоя внимания демонстрирует квадратичную зависимость от количества точек интереса.

С целью обработки более плотных облаков точек, обеспечивающих захват более мелких деталей объекта, используется кодировщик PointNet++ [5] ( $\mu_{enc}$ ) для получения двух наборов подвыборочных точечных признаков:  $f_1 = \mu_{enc}(P1)$  и  $f_2 = \mu_{enc}(P2)$ , где  $f_1, f_2 \in \mathbb{R}^{N' \times d_{sub}}$ ,  $N' < N$  — количество подвыборочных точек, а  $d_{sub}$  — размерность подвыборочных точечных признаков. К полученным подвыборочным точечным признакам применяется операция внимания на основе скалярного произведения.

$$Attn_{12} = \text{softmax}\left(\frac{f_1 f_2^T}{\sqrt{d_{sub}}}\right) f_2, \quad f_{12} = [f_1, Attn_{12}].$$

Объединенные подвыборочные точечные признаки  $f_{12} \in \mathbb{R}^{N' \times d_{sub}}$  представляют собой конкатенацию  $f_1$  и ре-

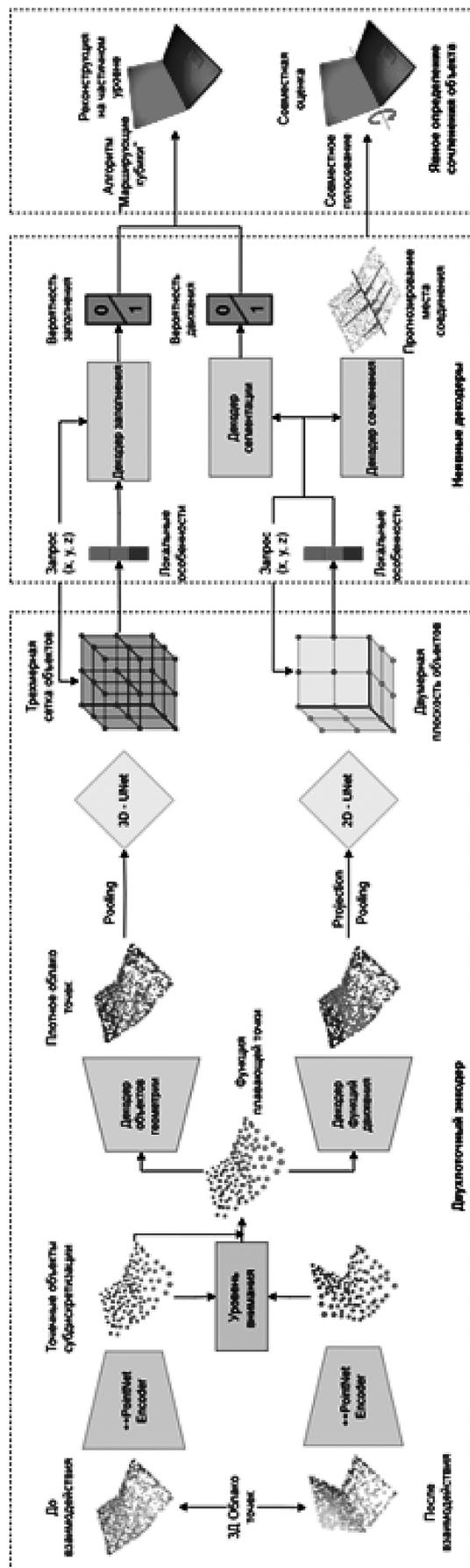


Рис. 3. Архитектура модели Ditto

зультата операции внимания. Далее, используются два декодера PointNet++ ( $v_{geo}$  и  $v_{art}$ ) для преобразования объединенных подвыборочных точечных признаков в плотные признаки, выровненные с исходными точками.

$$f_{geo} = v_{geo}(f_{12}) \quad \text{and} \quad f_{art} = v_{art}(f_{12}),$$

где  $f_{geo}, f_{art} \in \mathbb{R}^{N \times d_{dense}}$  — это  $d_{dense}$ -мерные точечные признаки, выровненные с P1. Использование двух отдельных наборов плотных точечных признаков обусловлено тем, что реконструкция геометрии в основном использует статичные наблюдения, в то время как оценка параметров сочленения в большей степени опирается на соответствие между двумя наблюдениями. Кроме того, обработка этих признаков осуществляется отдельно.  $f_{art}$  проецируется на двумерные плоскости признаков, а  $f_{geo}$  — на воксельные сетки, аналогично ConvONets [6]. Точки, попадающие в одну и ту же ячейку пикселя или вокселя, агрегируются с помощью max pooling. Данная операция проецирования существенно снижает вычислительные затраты, сохраняя пространственное распределение точечных признаков. Проецирование применяется к трем каноническим плоскостям  $s$  и грубой воксельной сетке  $v$ . Полученные плоскости и сетка признаков обрабатываются независимыми 2D и 3D UNet [7]. Выходная воксельная сетка признаков используется для неявного декодера геометрии, а плоскости признаков — для неявных декодеров параметров сочленения. Неявный декодер геометрии требует плотные сетки признаков для точной локальной обработки, в то время как для оценки параметров сочленения достаточно разреженных плоскостей признаков.

### Неявные декодеры

Опираясь на результаты современных исследований [8, 9, 10], демонстрирующих непрерывность и универсальность неявных нейронных представлений, разрабатываются неявные декодеры для задач восстановления геометрии и определения параметров сочленения. Поскольку обе задачи требуют детального анализа геометрии, входными данными для неявных декодеров являются локальные признаки. Локальные признаки вычисляются на основе сетки/плоскостей признаков с использованием трилинейной/билинейной интерполяции по заданной трехмерной координате запроса  $p \in \mathbb{R}^3$ .

Предлагается неявный декодер геометрии, реализующий отображение координаты  $p \in \mathbb{R}^3$  в вероятность заполнения  $o(p)$  в данной координате. Значение  $o(p)$  должно быть равно 1, если точка  $p$  находится внутри объекта, и 0 в противном случае. Локальный признак  $\psi_p^v$  извлекается из сетки признаков  $v$  посредством трилинейной интерполяции. С учетом координаты запроса  $p$  и локального признака  $\psi_p^v$  неявный декодер геометрии предсказывает вероятность заполнения:

$$f_{0_o}(p, \psi_p^v) \rightarrow o(p).$$

Неявные декодеры сочленения отображают произвольную точку  $p$ , расположенную внутри объекта, метку сегментации и параметры шарниров, относящиеся к данной точке. Рассматривается только внутреннее пространство объекта, поскольку оценка параметров сочленения имеет смысл только для точек внутри него. Локальный признак  $\psi_{pm}^c$  извлекается из плоскостей признаков с помощью билинейной интерполяции.

Сегментация. В рамках предположения, что в результате взаимодействия изменяется только состояние одного шарнира, предлагается сегментировать объект на статическую и подвижную части для каждого взаимодействия. Следовательно, прогнозируется бинарная метка сегментации  $s(p_{in})$ , где 0 является статической частью, а 1 — подвижная часть. Неявный декодер сегментации предсказывает вероятность сегментации, основываясь на локальном признаке объекта:

$$f_{0_{seg}}(p_{in}, \psi_{p_{in}}^c) \rightarrow s(p_{in}).$$

Параметры сочленения. Несмотря на то, что параметры сочленения являются глобальными характеристиками сочлененного объекта, для повышения эффективности структурированного представления признаков и повышения надежности оценки за счет процедуры голосования используется представление для каждой точки. Плоскости признаков используются совместно для прогнозирования параметров сочленения и сегментации, поскольку параметры сочленения могут быть выведены на основе таких признаков движения, как сегментация. Первым шагом является применение неявного декодера для определения (предсказания) типа шарнира  $P_{jtype}$ .

$$f_{0_{type}}(p_{in}, \psi_{p_{in}}^c) \rightarrow P_{jtype}(p_{in}).$$

Далее используются два неявных декодера которые служат для определения параметров и состояний призматических и вращательных шарниров. Призматический шарнир определяется направлением оси трансляции, представленным единичным вектором  $u^p$ . Состояние шарнира определяется расстоянием трансляции  $c^p$ , возникающим в результате взаимодействия. Параметры вращательного шарнира включают направление оси вращения  $u^r$ . В отличие от призматического шарнира, положение оси вращательного шарнира также является значимым параметром. В соответствии с подходом Li et al [11], положение шарнира относительно точки  $p_{in}$  определяется как проекция  $p_{in}$  на ось, которая представлена единичным вектором  $d_{p_{in}}^r$  для направления проекции и скаляром  $h_{p_{in}}^r$  для расстояния проекции. Со-

стояние вращательного шарнира определяется углом поворота  $c^r$ , возникающим в результате взаимодействия. Непосредственное предсказание этих параметров осуществляется с помощью неявных декодеров:

$$f_{\theta_{param_p}}(p_{in}, \psi_{p_{in}}^c) \rightarrow [u^p, c^p], \lambda$$

$$f_{\theta_{param_r}}(p_{in}, \psi_{p_{in}}^c) \rightarrow [u^r, d_{p_{in}}^r, h_{p_{in}}^r, c^r]$$

Так, немного подкорректировав данные формулы можно адаптировать и для сканирования обычных объектов. Вход: Облако точек  $P \in R^{(N \times 3)}$  (после возможной регистрации нескольких сканов)

1. Кодирование геометрии:

$$f_{geo} = \mu_{enc}(P) \in R^{M \times d_{geo}},$$

где  $\mu_{enc}$  — оптимизированный кодировщик PointNet++ или другая архитектура кодирования.

$M$  — количество точечных признаков

$d_{geo}$  — размерность точечных признаков

2. Проекция признаков:

Используем только проецирование на воксельную сетку:

$$v = Projection(f_{geo})$$

3. Обработка признаков:

$$v_{unet} = 3D_{UNet}(v)$$

4. Неявный декодер геометрии:

Для произвольной точки  $p \in R^3$ :

$\psi_p^v = \text{TrilinearSampling}(v_{unet}, p)$  — Извлечение локальных признаков из воксельной сетки

$o(p) = f_{\theta_o}(p, \psi_p^v)$  — Предсказание вероятности занятости

Вывод: 3D-модель объекта, полученная путем извлечения из поверхности из неявного поля занятости  $o(p)$ . При реализации и обучении нейронной сети можно получить следующий результат, указанный на рисунке 4.

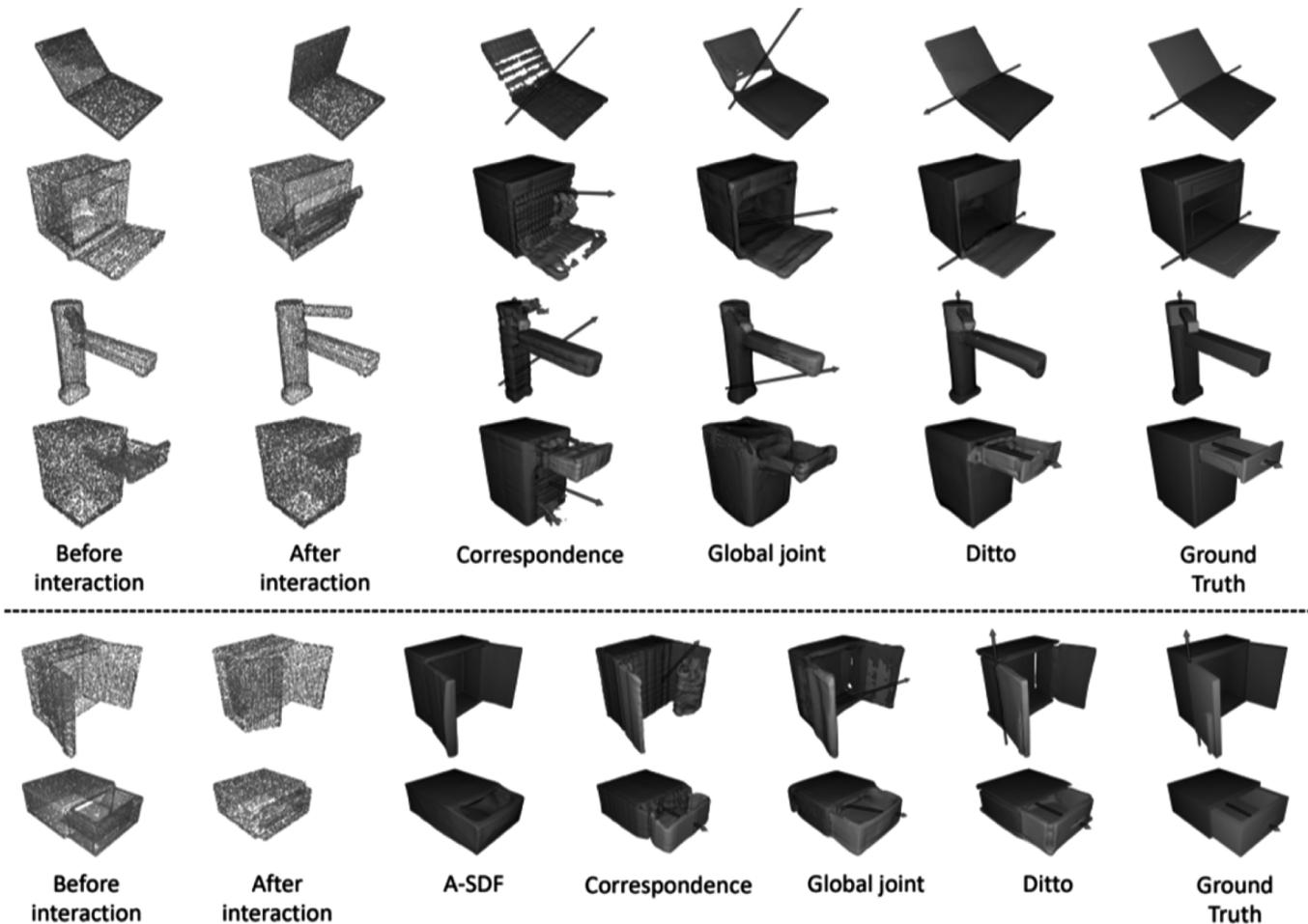


Рис. 4. Представлены результаты реконструкции ранее не наблюдаемых сочлененных объектов [12]

**Улучшение работы путем интеграции с алгоритмами регистрации облаков точек**

Использование алгоритмов регистрации облаков точек в трехмерном пространстве, таких как Iterative Closest Point (ICP) [13], может значительно улучшить процесс создания цифровых двойников шарнирно-сочлененных объектов, особенно в случаях, когда доступно несколько изображений объекта, частично охватывающих его поверхность.

Алгоритм Ditto использует информацию о взаимодействии с объектом для установления соответствий между двумя облаками точек до и после манипуляций над ним. Однако, если на вход поступают множественные, не выровненные неполные изображения, эффективность Ditto может начать снижаться. Справиться с данной проблемой возможно, за счет использования алгоритма регистрации облаков точек, который призван выровнять и объединить несколько частичных изображений в один полный и когерентный набор облаков точек. Это обеспечит более полную и точную геометрическую информацию для Ditto, что улучшает процесс реконструкции и оценки его сочленений.

Для использования ICP в связке с Ditto, необходимо, верно, выстроить интеграцию. Перед подачей данных в Ditto, необходимо выполнить этап предварительной обработки, а именно, получить как можно больше изображений объекта, максимально охватывающих все поверхность. Выделить из изображений облака точек с последующим удалением шумов и выбросов, с целью улучшения входных данных. После подготовки начинается использование алгоритма ICP для выравнивания и объединения облаков точек. Процесс начинается с выбора одного облака точек в качестве опорного (целевого). Затем происходит итеративное выравнивание остальных облаков точек относительно опорного, минимизируя расстояние между соответствующими точками. На этом этапе также происходит настройка параметров ICP, таких как максимальное количество итераций, порог сходимости, и выбор начального приближения.

После выравнивания и объединения всех облаков точек в одно полное облако происходит подача данных в Ditto. В случае, если Ditto требует два облака точек (до и после взаимодействия), необходимо либо создать эти облака искусственно (например, путем небольшого смещения частей объекта в виртуальном пространстве), либо адаптировать Ditto для работы с одним облаком точек. На рисунке 5 представлена схема применения алгоритма ICP в процессе создания цифровых двойников.

**Алгоритм Iterative Closest Point (ICP)**

Целью алгоритма ICP является нахождение оптимального преобразования, состоящего из вращения и транс-



Рис. 5. Схема применения алгоритма ICP в процессе создания цифровых двойников

ляции, для выравнивания и сопоставления между собой двух облаков точек. Данные облака точек получаются за счет большего количества изображений одного объ-

екта, сделанных с разных ракурсов, максимально охватывающих всю поверхность объекта.

Перед тем как рассмотреть принцип действия алгоритма, рассмотрим основные параметры и их обозначения:

- $P$  — опорное облако точек (целевое), представленное множеством точек  $P = \{p_i\}_{i=1}^{N_p}$ , где  $p_i \in R^3$  — координата  $i$ -й точки, а  $N_p$  — общее количество точек в опорном облаке;
- $Q$  — входное облако точек (исходное), представленное множеством точек  $Q = \{q_j\}_{j=1}^{N_Q}$ , где  $q_j \in R^3$  — координата  $j$ -й точки, а  $N_Q$  — общее количество точек в исходном облаке;
- $\epsilon$  — порог сходимости, определяющий критерий остановки алгоритма на основе изменения функции ошибки;
- $t_{\max}$  — максимальное количество повторений алгоритма.
- $T_0$  — начальное преобразование (используется вариативно), представляющее собой первичную оценку от выравнивания облаков точек;
- $T^*$  — оптимальное преобразование, состоящее из матрицы вращения  $R$  и вектора трансляции  $t$ , минимизирующее расстояние между соответствующими точками в облаках точек  $P$  и  $Q$ .

Старт алгоритма начинается с его инициализации, которая начинается с установки начального преобразования  $T_k$  равным  $T_0$  (если используется), в обратном случае установить  $T$  равным единичной матрице. Затем следует инициализировать счетчик итераций  $k$  значением 0. Далее следует перейти к самому процессу, а именно к повторению шагов, указанные ниже, пока  $k < t_{\max}$ . После начинается нахождение соответствий для каждой точки  $q_j \in Q$ , преобразование точки  $q'_j = T \cdot q_j$  и нахождение ближайшей точки  $p_j \in P$ , минимизирующую евклидово расстояние:

$$p_j = \operatorname{argmin}_{p_i \in P} \| q'_j - p_i \|^2$$

Затем идет вычисление преобразования  $T_{k+1}$  (вращение  $R$  и трансляция  $t$ ), минимизирующее функцию ошибки:

$$E(T) = \frac{1}{N_Q} \sum_{j=1}^{N_Q} |T \cdot q_j - p_j|^2$$

Задача минимизации решается, как правило, методом наименьших квадратов. Для вычисления матрицы вращения и вектора трансляции может быть использован алгоритм Хорна.

Так следует применить преобразование и обновить преобразование:  $T \leftarrow T_{k+1}$ . И провести проверку сходимости, если изменение функции ошибки на текущей итерации меньше порога сходимости:

$$\Delta E = |E(T_{k+1}) - E(T)| < \epsilon$$

В таком случае алгоритм завершает работу. После идет увеличение счетчика итераций:  $k \leftarrow k+1$

Следует перейти так же к объяснению некоторых технических деталей и преимуществ данного алгоритма в работе, связанных с применением данного способа для конечного понимания общих принципов работы и необходимости его дополнения к основной формуле. Во-первых, выбор алгоритма регистрации ICP является одним из наиболее распространенных алгоритмов регистрации, но существуют и другие, более продвинутые алгоритмы, такие как Feature-based регистрация, Normal Distributions Transform (NDT), или Variational ICP, которые могут быть подходящими в зависимости от характеристик данных сканирования (например, наличие текстуры, количество шумов, степень перекрытия между сканами). Во-вторых, более полные и точные данные о геометрии объекта, получаемые после регистрации, позволяют Ditto более точно реконструировать 3D-модель. Так же применение данного алгоритма представляет более точное представление геометрии частей объекта способствует более надежной оценке параметров сочленения. Способствует автоматизация процесса регистрации позволяет сократить необходимость в ручной обработке данных сканирования, что упрощает и ускоряет процесс создания цифровых двойников.

#### ЛИТЕРАТУРА

1. 3D ML. Часть 6: Обзор алгоритмов семантической сегментации облака точек [Электронный ресурс]. URL: <https://habr.com/ru/companies/itmai/articles/534036/> (дата обращения: 01.03.2025).
2. Charles R. Qi., Li Yi., Hao Su., Leonidas J. Guibas. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space // 2017.
3. Xiaolong Li., He Wang, Li Yi., Leonidas Guibas, A. Lynn Abbott, Shuran Song. Category-Level Articulated Object Pose Estimation // 2019.
4. Jiapeng Tang, Jiabao Lei, Dan Xu, Feiyang Ma, Kui Jia, Lei Zhang. SA-ConvONet: Sign-Agnostic Optimization of Convolutional Occupancy Networks // 2021.
5. Трехмерное распознавание: текущее состояние и тенденции // Автоматика и телемеханика. 2022. № 4
6. Songyou Peng, Michael Niemeyer, Lars Mescheder, Marc Pollefeys, Andreas Geiger. Convolutional Occupancy Networks // 2020.
7. Шелковников Евгений Юрьевич, Шляхтин Константин Алексеевич, Шелковникова Татьяна Евгеньевна, Егоров Станислав Феликсович применение нейронной сети архитектуры U-Net для сегментации СТМ-изображений // Химическая физика и мезоскопия. 2019. №2. URL: <https://cyberleninka.ru/article/n/primenenie-nei-ronnoi-seti-arhitektury-u-net-dlya-segmentatsii-stm-izobrazhenii> (дата обращения: 02.03.2025).

8. Pete Florence, Corey Lynch, Andy Zeng, Oscar Ramirez, Ayzaan Wahid, Laura Downs, Adrian Wong, Johnny Lee, Igor Mordatch, Jonathan Tompson. Implicit Behavioral Cloning // 2021.
9. Zhenyu Jiang, Yifeng Zhu, Maxwell Svetlik, Kuan Fang, Yuke Zhu. Synergies Between Affordance and Geometry: 6-DoF Grasp Detection via Implicit Representations // 2021.
10. Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin. Occupancy Networks: Learning 3D Reconstruction in Function Space Long Beach, CA, USA: IEEE, 2019. С. 4455–4465.
11. Xiaolong Li, He Wang, Li Yi, Leonidas J Guibas, A Lynn Abbott, and Shuran Song. Category-level articulated object pose estimation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 3706–3715, 2020
12. Jiang Z., Hsu C.-C., Zhu Y. Ditto: Building Digital Twins of Articulated Objects from Interaction // 2022.
13. Якимчук А.В. Разработка системы технического зрения для определения позиции и ориентации объекта на основе 3D-модели // StudNet. 2020. №10. URL: <https://cyberleninka.ru/article/n/razrabotka-sistemy-tehnicheskogo-zreniya-dlya-opredeleniya-pozitsii-i-orientatsii-obekta-na-osnove-3d-modeli> (дата обращения: 01.03.2025).

---

© Тимаков Кирилл Александрович (Tim1997.10@yandex.ru)  
Журнал «Современная наука: актуальные проблемы теории и практики»