

Управление в технических системах

© 2025 г. ЯССЕР ДЕРРАП (derrar.yasser@gmail.com),

ФАРАХ САИДИ (saidifarah.mimouni@gmail.com)

(LAT, Лаборатория автоматика, Университет Тлемсена, Тлемсен, 13000, Алжир),

АБЕД МАЛЬТИ (abed.malti@gmail.com)

(GBM, Лаборатория биомедицины, Университет Тлемсена, Тлемсен, 13000, Алжир)

ВИЗУАЛЬНОЕ СЕРВОУПРАВЛЕНИЕ ДЛЯ ДЕФОРМИРУЕМЫХ ОБЪЕКТОВ С ИСПОЛЬЗОВАНИЕМ ГЕОМЕТРИЧЕСКИХ ПРИМИТИВОВ И ЗАРАНЕЕ СПЛАНИРОВАННЫХ ТРАЕКТОРИЙ

Представлена новая техника жесткого выравнивания перспективной камеры с деформациями нежесткого объекта, использующая адаптивный подход визуального сервопривода. В отличие от существующих методов, предлагаемый подход не опирается на какие-либо параметрические или модельные априорные данные о деформации. Предполагая наличие заранее спланированной траектории камеры, наблюдающей за слитно жестким объектом, метод направлен на выравнивание этой траектории во время фазы выполнения, используя только наиболее релевантные ориентиры в качестве априорной информации. Подход не зависит от каких-либо параметрических или непараметрических моделей физики деформации. Он формулируется как задача отслеживания, встроенная в оптимальную схему визуального управления. Этот процесс отслеживания включает в себя визуальный сервопривод геометрических особенностей деформируемого объекта, соединяя этапы планирования и выполнения. Оптимальное визуальное управление определяется с использованием критерия взвешенного метода наименьших квадратов, который минимизирует расстояние между опорными признаками и наблюдаемыми в режиме реального времени. Веса представляют собой зависящие от времени плавные функции, которые кодируют значимость видимых особенностей объекта. Экспериментальные результаты демонстрируют способность метода адаптироваться к различным заранее спланированным траекториям и типам деформаций без необходимости предварительных знаний, а также его устойчивость к шуму при обнаружении признаков на изображении.

Ключевые слова: визуальное обслуживание, деформирование объектов, планирование движения, отслеживание траектории.

DOI: 10.31857/S0005231025020058, EDN: IQPGLO

1. Введение

Визуальное сервоприводное управление относится к использованию данных компьютерного зрения, полученных с одной или нескольких камер, для управления их движением. В отличие от задач отслеживания на основе зрения с жесткими объектами, которые достигли определенной зрелости, отслеживание с помощью визуального сервопривода на неупругих объектах

остается сложной задачей. В последние годы это вызвало большой интерес в сообществах компьютерного зрения и робототехники [1]. Многочисленные потенциальные применения нацелены на такие области, как дополненная реальность, медицинская визуализация, роботизированные манипуляции, и включают работу с огромным разнообразием объектов: ткани, бумага, резина, вязкие жидкости, кабели, пищевые продукты, органы и т.д [2]. Одна из характеристик отслеживания деформируемых объектов заключается в том, что форма объекта изменяется в процессе отслеживания. Недавние методы предложили гибридное визуальное сервоприводное управление, разделяющее поступательную скорость по оси Z и три угловые скорости. Такой метод улучшил производительность классического HVS (гибридного визуального сервоприводного управления) как в плоскости изображения, так и в рабочем пространстве [3]. Однако разделение степеней свободы камеры накладывает ограничения на коррекцию движения камеры и может привести к неравномерному отслеживанию деформации объекта. Отслеживание геометрии деформируемых объектов, таких как веревка и ткань, затруднено из-за непрерывной природы объекта (т.е. бесконечного числа степеней свободы) [4, 5]. Большинство современных методов отслеживания с помощью визуального сервопривода требуют как знания физических свойств объекта, так и мониторинга деформаций в реальном времени. Несколько исследований сосредоточены на мониторинге деформаций в реальном времени, но не на одновременном управлении деформацией [6]. Эти методы являются высокопараметрическими и требуют специфической тонкой настройки, которая сильно зависит от физики деформаций. В этой работе предлагается надежный подход к отслеживанию деформации объекта с помощью оптимального визуального сервопривода. Предполагается наличие заранее спланированной эталонной траектории камеры, наблюдающей за целевым объектом. Предполагается, что этот объект описан в сетке, состоящей из набора треугольных геометрических примитивов. Не предполагается никакого предварительного знания или информации о типе или физике деформации. В качестве эталонного деформируемого объекта использована печень. Задача формулируется следующим образом: пусть камера следует по эталонной траектории и наблюдает за печенью. Если печень деформируется, как лучше всего адаптировать траекторию камеры? В контексте роботизированной лапароскопии это стандартная проблема, которая обычно возлагается на оператора, управляющего камерой (лапароскопом) с помощью джойстика по рекомендации хирурга. Эта задача может быть полностью автоматизирована при решении двух основных вопросов: во-первых, необходима соответствующая система отслеживания печени, позволяющая следить за 2D-признаками на печени. Во-вторых, необходим критерий автоматизации, позволяющий адаптироваться к деформациям и соответственно изменять траекторию камеры. В данной статье предлагается новая стратегия для решения второй проблемы.

Вклад данной работы иллюстрируется на рис. 1. Можно подвести итоги следующим образом.



Рис. 1. Блок-схема алгоритма визуального сервопривода на основе треугольных сеток для деформируемых объектов.

Эта диаграмма иллюстрирует процесс работы алгоритма визуального сервопривода на основе треугольных сеток, используемого для управления движением камеры при наблюдении деформируемого объекта. Процесс начинается с инициализации эталонной траектории и данных о сетке, после чего происходит итерация по каждой точке на траектории. Для каждой точки алгоритм обрабатывает каждый треугольник в сетке объекта, вычисляя ключевые параметры, такие как веса и матрицы взаимодействия. Затем визуальный сервоприводный цикл итеративно корректирует позицию и ориентацию камеры, чтобы минимизировать ошибку между текущим изображением деформированного объекта и эталонным изображением (функция задачи), в конечном итоге создавая оптимизированную траекторию, учитывающую деформацию объекта.

- 1) Определение планирования движения на основе камеры для компенсации деформаций нежестких объектов жестким движением камеры. Это определение позволяет формализовать визуальное взаимодействие между деформируемым объектом и камерой с помощью набора примитивов «камера-треугольник». Эти треугольники являются геометрическими элементами сетки объекта.
- 2) Если связать функцию непрерывного веса с каждой парой примитивов камера-треугольник, то непрерывность вдоль траектории камеры и позволяет обеспечить плавное отслеживание.

- 3) Решение задачи визуального обслуживания и отслеживания в оптимальной структуре визуального контроля. Этот предложенный метод не предполагает каких-либо априорных данных о деформации объекта. Вышеуказанные весовые функции кодируют релевантность примитивов камера-треугольник. Наиболее важные примитивы, связанные с мишенями, имеют больший вес и позволяют найти оптимальное движение камеры для компенсации нежестких деформаций.
- 4) Экспериментальная валидация демонстрирует надежность подхода относительно шумов в обнаружении изображений, вариабельности траекторий и типов деформаций.

Статья организована следующим образом. В разделе 2 представлены связанные современные работы. В разделе 3 вводятся основные инструменты для моделирования. В разделе 4 описывается визуальное сервоприводное управление с оптимальным визуальным контролем и отслеживанием. В разделе 5 представлены экспериментальные результаты и комментарии. Раздел 6 подводит итоги статьи и очерчивает направления будущих работ.

2. Связанная работа

Отслеживание и визуальное сервоприводное управление неупругими объектами – это недавно возникшие открытые проблемы. Многие подходы предлагают использовать физические или основанные на данных механические модели, встроенные в классические алгоритмы. Эти методы являются высокопараметрическими и тонко настраиваются для каждого конкретного типа деформируемых объектов. В этой статье предлагается универсальный подход, который позволяет жесткому визуальному сервоприводу камеры максимально соответствовать визуальным признакам объектов, подверженных неупругой деформации. В этом разделе рассматривается литература по визуальному сервоприводу, визуальному отслеживанию и выравниванию изображений.

2.1. Визуальное обслуживание жестких объектов

Классическое визуальное сервоприводное управление – это метод управления движением робота с использованием обратной связи в реальном времени от датчиков зрения [7–9]. Визуальное сервоприводное управление имеет два основных подхода: позиционное визуальное сервоприводное управление и управление, основанное на изображениях [10, 11]. В этой статье используется позиционный визуальный контроль, так как предполагается отслеживание функций $2D$ на целевом деформируемом объекте. Отношения между камерой и объектом представлены матрицей взаимодействия. Эта матрица может быть определена в пространстве изображения точками, линиями или эллипсами и моментами [12, 13]. В этой работе используется матрица взаимодействия на основе точек. Она используется в качестве элементарного

строительного блока для матриц взаимодействия на основе треугольников. Треугольники являются примитивами, образующими поверхность деформируемой формы. Поскольку эта проблема существует уже давно, перечень визуального обслуживания жестких объектов является обширным и не может быть охвачен в настоящем разделе. Здесь рассмотрим некоторые основные результаты этих работ. Джанаби и др. [14] используют уменьшенный набор отверстий, окружностей и клиньев как характеристики для их доступности во многих промышленных частях и для их легкого и надежного извлечения. В [15] моменты, рассчитанные с помощью сегментации изображения, используются для определения аналитической формы матрицы взаимодействия. В [16] предложена техника визуального сервопривода на основе маркеров для автоматического позиционирования камеры в случае роботизированной минимально инвазивной хирургии. Такой подход заключается в том, чтобы держать последующий хирургический инструмент в пределах изображения камеры. Это делается путем постоянной настройки позы эндоскопической манипуляции с камерой вместо того, чтобы держать инструмент в центре изображения. Кроме того, матрица взаимодействия может быть определена гибридным визуальным сервированием, совмещающим визуальное сервирование на основе изображений и позиционное визуальное сервирование. В [17] предложена стратегия управления в реальном времени и без инверсии с помощью алгоритма модельного прогнозирования на основе выборки для визуальных методов, основанных как на изображениях, так и на положении. Они учитывали системные ограничения и параметры неопределенности, связанные с роботом и измерениями камеры. В данной статье не рассматриваются неопределенности модели и не используется схема предиктивного управления. Вместо этого используется структура оптимального визуального управления, чтобы компенсировать как шум, так и деформации. Предлагается использовать непрерывные весовые функции на треугольных примитивах для учета их значимости при отслеживании опорной траектории.

2.2. Визуальное обслуживание нежестких объектов

Роботизированная манипуляция нежесткими объектами является сложной задачей из-за деформаций, которые возникают и добавляют несколько степеней свободы к исходной задаче с жесткими объектами. Среди деформаций можно упомянуть сдвиг, масштабирование, растяжение, кручение, сжатие и т.д. Человеческие органы являются актуальными примерами нежестких объектов с разнообразными формами и деформационными поведением [18]. Использование таких объектов с визуальным сервированием позволяет автоматизировать проблемы в медицинском и хирургическом контекстах. Например, дополненная реальность накладывает МРТ (магнитно-резонансную томографию) изображения на поток живой лапароскопии [2], роботизированная лапароскопия – для направления лапароскопа и т.д. В [19] представлен контроллер, основанный на глубокой нейронной сети для управления положением и формой деформируемых объектов с неизвестными деформационны-

ми свойствами. Рассмотрены нелинейные свойства деформируемых объектов и использована многослойная нейронная сеть для моделирования функции отображения. В этой работе не используется никакая предварительная модель деформации органа, а используется набор функций непрерывного веса, который позволяет учитывать соответствующие регионы, чтобы сосредоточиться на нем. В [20] представлена гистограмма ориентированных морщин для описания вариаций формы сильно деформируемого объекта. Характеристики деформируемого объекта рассчитываются путем применения фильтров Габора и выделения компонентов высокой и низкой частоты, предварительно была вычислена визуальная обратная связь с использованием фазы обучения в автономном режиме, которая хранит соответствие между этими визуальными характеристиками и скоростью конечного эффектора. Предлагаемый здесь подход использует оптимальную настройку визуального управления, которая осуществляет встроенную коррекцию камеры без необходимости получения данных или тренировки нейронных сетей. В [21] разработан сервоалгоритм, который может изучать нелинейную деформационную функцию вместе с процессом манипуляции. При этом используется гауссовская регрессия процессов для моделирования и изучения параметров деформации мягкого объекта. В данной статье предлагается общий подход, который позволяет жесткому видеосервису максимально адаптировать визуальные особенности объектов, подверженных нежесткой деформации. В предлагаемом способе используется треугольное сетчатое представление визуализируемого объекта, вычисляется относительное жесткое изменение положения камеры, которое соответствует наиболее подходящим примитивам треугольника, не полностью игнорируя остатки.

2.3. Визуальное отслеживание нежестких объектов

Отслеживающая деформация изучалась в нескольких работах [22–25]. В [26] представлен двухэтапный метод отслеживания объектов. Здесь использовался метод на основе ядра для эффективного обнаружения объекта в сложных условиях с перемещением камеры. Для повышения точности отслеживания использовался контурный метод, позволяющий четко следовать контуру объекта после определения его местоположения. В [27] предполагалось, что расширение объекта деформируется по сравнению с эталоном путем перемещения определенных контрольных точек последнего к контрольным точкам старого объекта. В предлагаемом здесь методе не рассматривается ни одной модели деформированного объекта, компенсация движения основана на соответствующих различиях характеристик изображения. В [28] представлена генеративная модель деформации тела, которая имеет способность выразить движение каждой основной части объекта. В [29] представлен новый подход для отслеживания деформируемой анатомической мишени в объемах 3D-ультразвука. Этот метод способен оценить деформации, вызванные физиологическими движениями пациента. Оценка смещения движущихся структур осуществляется с помощью подхода, основанного на интенсив-

ности, в сочетании с физической моделью. Это позволяет сделать оценку менее чувствительной к шуму на изображении. Предлагаемый здесь метод не обрабатывает 3D-реконструкцию объекта и фокусируется только на наилучшем движении камеры для компенсации деформаций и, в конечном счете, шума. В [30] был предложен метод 3D-слежения, который упорядочивается статистической моделью движения, полученной из биомеханического моделирования. Однако этот метод требует ручного определения конкретных точек целевого объекта (здесь предстательная железа) на каждом ультразвуковом кадре, чтобы управлять моделью. В [31] разработана система, которая отслеживает траекторию линии, нарисованной на гибком объекте. Также была оценена эффективность слежения со скоростью и точностью при наличии многих неопределенностей, основанных на динамической компенсации. В предлагаемом подходе нет шаблона, который нарисован на целевом объекте. Вместо этого набор весовых функций заранее определен для предварительно выбранных областей деформируемых объектов. Эти области определяются в наборе треугольников сетки объекта.

2.4. Выравнивание изображения

Выравнивание изображений заключается в выравнивании двух или более изображений одной и той же сцены по общей пространственной оси. Оно играет важную роль в компьютерном зрении и графике, таких как восстановление структуры из движения, 3D реконструкция, отслеживание движения и восстановление [32, 33]. Проблему, которая здесь решается, можно рассматривать как выравнивание изображения деформируемой сцены. Действительно, компенсация движения камеры, которая рассчитывается предлагаемым методом, также может быть использована для выравнивания текущего изображения с эталонным изображением по заданной траектории. В [34] предложен метод выравнивания края гибкого листового объекта по заданному отрезку линии в трехмерном пространстве. Произведена онлайн-оценка относительной позиции между концом края и рукой робота в визуальном законе управления сервированием. Метод авторов опирается на характерные точки, а не на особенности линий. В [35] фокусируются на параметрическом и непараметрическом методе выравнивания, которые имеют дополнительную силу. Предложено параметрическое выравнивание на основе признаков с использованием одной или нескольких гомографий, за которыми следует непараметрическое выравнивание в пиксельной форме. Метод авторов непараметричен относительно деформации объекта. В [36] предложен спектрально-пространственный взвешенный метод встроенного выравнивания распределения ядерного многообразия для классификации изображений дистанционного зондирования. Использован фильтр для выражения среднего спектра соседних образцов пикселей из каждого образца пикселя. В этой работе авторы рассматривают точки двумерных признаков как известные. В [37] предложен способ визуального сервопривода для перемещения ультразвукового зонда, используя робота для выравнивания плоскости изображения зонда с

иглой. Способ сегментирует иглу и обновляет набор визуальных признаков на основе модели иглы. Для отслеживания процесса центровки используется конечный автомат, а для управления датчиком в различных состояниях используются различные визуальные элементы. В этой статье авторы рассматривают перспективную модель камеры. Предложенный способ может быть применен к ультразвуковым датчикам, однако потребуется изменить модель проекции, которая используется в данной статье. В [38] была представлена система тактильной обратной связи на основе зрения, предназначенная для помощи в перемещении эндоскопического устройства во время капсульной эндоскопии. Она позволяет пользователю контролировать движение капсулы вдоль сгенерированного пути. Тактильный модуль также помогает оператору, преобразуя 3D-карты и относительные пути в направляющую виртуальную силу. Измеряя текущее относительное расстояние между вводом пользователя и границами карт, модуль тактильного управления проверяет, удаляется ли пользователь или приближается к стенкам толстой кишки, и генерирует обратную силу, чтобы помочь оператору во время навигационной процедуры. Пользователь также будет ощущать притягивающую виртуальную силу обратной связи в направлении сгенерированного пути, что поможет ему в навигации. Предложенный здесь метод может быть применен для подачи необходимой коррекции в тактильный модуль. Вместо компенсации траектории камеры корректируется траектория капсулы с помощью управления тактильным джойстиком.

3. Визуальные примитивы и интерактивное моделирование

3.1. Определения

3.1.1. Треугольник

Рассмотрим треугольник T как базовый геометрический примитив, состоящий из трех точек в трехмерной рабочей области. Пусть $L = \mathbf{R}^9$ – конфигурационное пространство этого примитива. Обозначим через $l = (X_1, X_2, X_3, Y_1, Y_2, Y_3, Z_1, Z_2, Z_3)^T$ конфигурацию T в L , где $(X_i, Y_i, Z_i)^T$, $1 \leq i \leq 3$ являются координатами трех вершин, образующих треугольник T . Индекс i представляет число вершин рассматриваемого треугольника.

3.1.2. Камера

Обозначим как C камеру, которая отображает один или несколько треугольных примитивов на 2D-объект в пространстве изображения. Пусть $C = \mathbf{SE}(3)$ будет конфигурационным пространством этого датчика. Обозначим через c конфигурацию камеры в пространстве C .

3.1.3 Чувство треугольника

Восприятие треугольника T камерой C можно охарактеризовать непрерывным отображением:

$$(1) \quad \Pi : C \times T \rightarrow I_{c,l},$$

$$(2) \quad (c, l) \mapsto \Pi(c, l).$$

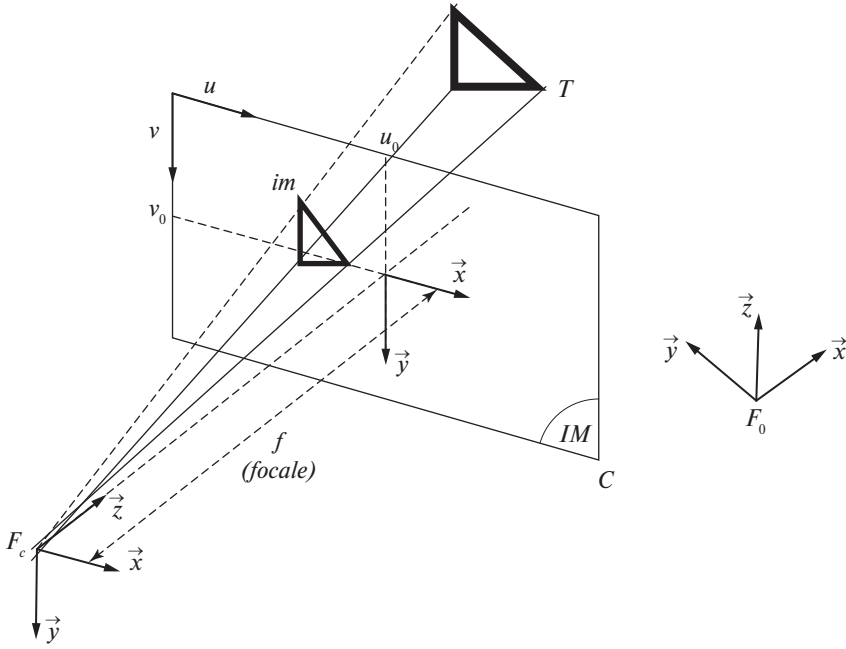


Рис. 2. Перспективная проекция треугольника на плоскость изображения камеры.

Это связывает с каждой конфигурацией камеры и треугольника функцию в пространстве изображения $I_{c,l} \subset \mathbf{R}^6$. Пространство изображения состоит из проекции тройки точек, описывающих интересующий треугольник (см. рис. 2). Π обычно определяется для подмножества $C \times T$ соответствующим конфигурациям пары камеры и треугольника, который должен находиться в поле зрения камеры. В случае перспективной камеры Π может быть записан как

$$(3) \quad \Pi(c, l) \left(\begin{array}{l} x_1 = X_1/Z_1, \quad x_2 = X_2/Z_2, \quad x_3 = X_3/Z_3, \\ y_1 = Y_1/Z_1, \quad y_2 = Y_2/Z_2, \quad y_3 = Y_3/Z_3 \end{array} \right)^\top.$$

3.2. Локализация и визуальное обслуживание

3.2.1. Уравнение локализации

Рассмотрим камеру в конфигурации $c \in C$. Давайте рассмотрим m треугольников T_1, \dots, T_m известных конфигураций $l_1, \dots, l_m \in L$, где каждый из них виден с камеры C . Таким образом, каждая пара (C, T_i) из треугольника-камеры описывается уравнением локализации:

$$(4) \quad \Pi(c, l_i) = im_i,$$

где $im_i \in I_{c,l_i}$ является проекцией T_i в пространстве изображений. $im_i = (x_1, x_2, x_3, y_1, y_2, y_3)^\top$, c и l_i рассматриваются в следующем разделе.

3.2.2. Визуальное обслуживание справочной конфигурации

Если ожидается, что камера будет следовать по эталонной траектории, просматривая эталонный объект, который может деформироваться, то результат уравнения (4), как ожидается, будет в районе эталонной конфигурации камеры и эталонной конфигурации объекта. Если пренебречь остатком линеаризации, то уравнение (4) может быть линеаризовано следующим образом:

$$(5) \quad \frac{\partial \Pi}{\partial c}(c_0, l_{0i})(c - c_0) + \frac{\partial \Pi}{\partial l}(c_0, l_{0i})(l_i - l_{0i}) = im_i - im_{0i}.$$

Здесь c_0 – это эталонная конфигурация камеры на заранее определенной траектории, l_{0i} – ожидаемая эталонная конфигурация рассматриваемого треугольника с эталонными характеристиками im_{0i} в пространстве изображений I_{c,l_i} . Во время выполнения обе эти эталонные конфигурации могут отличаться от конфигураций среды выполнения по двум основным причинам: дрейф камеры и деформация объекта. Эти соображения верны, если предположить, что локализация выполняется на высокой частоте [39] (быстрее, чем и контрольная петля камеры, и динамика деформации объекта).

im_i и im_{0i} соответственно представляют собой изображение треугольника T_i на датчике C и ожидаемое изображение, т.е. изображение, которое было бы видно из конфигурации c_0 , если не было дрейфа камеры и деформации объекта.

$\frac{\partial \Pi}{\partial c}$ – матрица Якобиана порядка 6 представляет изменение изображения по отношению к изменению конфигурации камеры. Эта матрица Якоби происходит от геометрических свойств камеры и треугольника и представляет собой так называемую матрицу взаимодействия в классической визуальной сервирующей терминологии. $\frac{\partial \Pi}{\partial l}$ – матрица Якоби порядка 6×9 , представляющая изменение изображения в зависимости от изменения конфигурации треугольника. Эта матрица Якоби является нулевой матрицей в случае недеформирующихся объектов. Большинство предыдущих работ смоделировали эту матрицу либо с помощью физических уравнений деформаций, либо с помощью алгоритмов машинного обучения, управляемых данными [40–42]. В этой работе используется оптимальная структура визуального контроля для внедрения контрольного цикла алгоритма визуального обслуживания. Этот подход позволяет учитывать деформационное поведение без необходимости численного вычисления или регресса деформаций. Сначала сосредоточимся на аналитической формулировке матрицы взаимодействия. Хорошо известно, что для одной точки в 3D-координатах (X, Y, Z) , воспринимаемой в 2D-координатах (x, y) , матрица взаимодействия, которая связывает вариации положения камеры с вариациями 2D-координат проецируемого изображения, записывается как [15]

$$(6) \quad L(x, y, Z) = \begin{bmatrix} L_u(x, y, Z) \\ L_v(x, y, Z) \end{bmatrix} = \begin{bmatrix} \frac{1}{Z} & 0 & \frac{-y}{Z} & -xy & 1 + x^2 & -y \\ 0 & \frac{1}{Z} & \frac{-x}{Z} & -(1 + y^2) & xy & x \end{bmatrix}.$$

Предполагается, что внутренние компоненты камеры известны и отличаются от координат пикселей. Если рассматривать треугольник T в координатах отсчета $l_0 = (X_1^0, X_2^0, X_3^0, Y_1^0, Y_2^0, Y_3^0, Z_1^0, Z_2^0, Z_3^0)^\top$ (см., например, рис. 2), то можно представить связанную матрицу взаимодействия как

$$(7) \quad \frac{\partial \Pi}{\partial c}(c_0, l_{0i}) = \begin{bmatrix} L_u(x_1^{0i}, y_1^{0i}, Z_1^{0i}) \\ L_u(x_2^{0i}, y_2^{0i}, Z_2^{0i}) \\ L_u(x_3^{0i}, y_3^{0i}, Z_3^{0i}) \\ L_v(x_1^{0i}, y_1^{0i}, Z_1^{0i}) \\ L_v(x_2^{0i}, y_2^{0i}, Z_2^{0i}) \\ L_v(x_3^{0i}, y_3^{0i}, Z_3^{0i}) \end{bmatrix},$$

где $im^0 = (x_1^0, x_2^0, x_3^0, y_1^0, y_2^0, y_3^0, z_1^0, z_2^0, z_3^0)^\top$ является контрольной проекцией треугольника на плоскость камеры (фокусное расстояние и центр камеры считаются известными и заданными). Для получения более подробной информации об определении c и $\frac{\partial \Pi}{\partial c}(c_0, l_{0i})$ см. Приложение.

3.2.3. Взвешенное трехстороннее визуальное обслуживание на жестких объектах

Если считаем объект жестким, то $l_i = l_{0i}$ для всех m треугольников, и уравнение 5 упрощается

$$(8) \quad \frac{\partial \Pi}{\partial c}(c_0, l_{0i})(c - c_0) = im_i - im_{0i}.$$

Классические визуальные сервирующие подходы на жестких объектах заключаются в решении для позы камеры c такого уравнения для всего множества треугольников. Это уравнение определяет и решает локальную задачу, которая заключается в подгонке текущего вида треугольника к ожидаемому. Можно также назначить положительные вещественные веса треугольникам объекта, которые имеют отношение к текущей задаче. Вышеприведенное уравнение оказывается

$$(9) \quad w_i \frac{\partial \Pi}{\partial c}(c_0, l_{0i})(c - c_0) = w_i(im_i - im_{0i}).$$

Здесь w_i – присвоенный вес треугольнику L_i . Рассматривая весь набор треугольников, составляющих объекты с соответствующим набором весов, построим следующую систему уравнений:

$$(10) \quad W(c - c_0) = IM - IM_0,$$

где

$$(11) \quad W = \begin{pmatrix} w_0 \frac{\partial \Pi}{\partial c}(c, l_0) \\ \vdots \\ w_m \frac{\partial \Pi}{\partial c}(c, l_m) \end{pmatrix},$$

$$(12) \quad IM = \begin{pmatrix} w_1 im_1 \\ \vdots \\ w_m im_m \end{pmatrix} \quad \text{и} \quad IM_0 = \begin{pmatrix} w_1 im_{01} \\ \vdots \\ w_m im_{0m} \end{pmatrix}.$$

Для решения уравнения (10) для положения камеры в случае жесткого объекта можно принять во внимание множество факторов. Например, шум при измерении положения треугольника на изображении, неточность формы трехмерного объекта, которая была предварительно отсканирована или смоделирована, и т.д. Если число треугольников слишком велико, система уравнений (10) не имеет точного решения. Наоборот, если система уравнений (10) не определена, существует бесконечное число возможных положений камеры, которые удовлетворяют ему. В этом случае для решения могут быть приняты такие соображения, как решение о минимальных нормах, как это было сделано в [39]. В следующем разделе опишем третий вклад авторов, который заключается во встраивании взвешенного треугольного визуального сервировочного формализма в оптимальную структуру визуального контроля для нежестких объектов. Такое встраивание позволяет рассчитать коррекцию для наложения на движение камеры при отслеживании опорной траектории на деформирующемся объекте. Предложенный подход не учитывает каких-либо предварительных знаний или параметрической модели, которые описывают деформацию целевого объекта.

4. Обслуживание на деформируемых объектах с помощью визуального элемента управления, отображаемого через точку с запятой

В трехмерной реконструкции деформирующихся объектов из монокулярных представлений доказано, что любая деформированная форма лежит на минимальной энергии растяжения/сжатия [43–46]. Здесь не будем стремиться выводить 3D форму из 2D вида, нужно только компенсировать деформацию движением камеры. Поэтому определяем такое движение как то, при котором энергия растяжения/сжатия минимальна. Используя уравнение (5), это сводится к минимизации деформации, видимой через перспективную проекцию, что можно формализовать следующим образом:

$$(13) \quad \hat{c} = \underset{\tilde{c} \in C}{\operatorname{argmin}} \frac{1}{2} \sum_{i=1}^m \left\| \frac{\partial \Pi}{\partial l}(c_0, l_{0i})(l_i - l_{0i}) \right\|_2^2,$$

$$\text{s.t.} \quad \frac{\partial \Pi}{\partial l}(c_0, l_{0i})(l_i - l_{0i}) + \frac{\partial \Pi}{\partial c}(c_0, l_{0i})(\tilde{c} - c_0) = (im_i - im_{0i}).$$

В работе некоторым пространственным деформациям придается большее значение, чем другим. Эта цель оформляется через ассоциацию действительных положительных значений веса и треугольных примитивов. Таким обра-

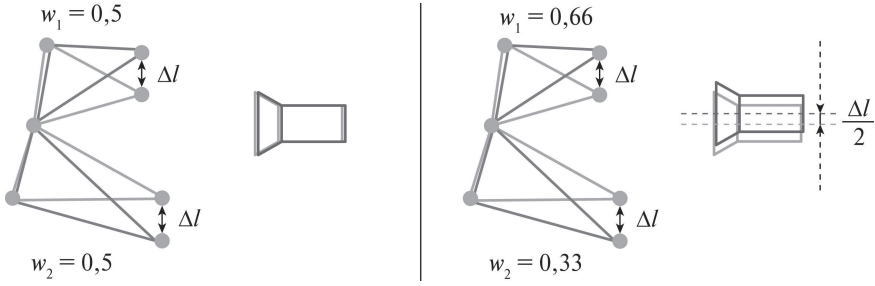


Рис. 3. Иллюстрация весового эффекта.

зом, вышеуказанный критерий может быть переписан как

$$(14) \quad \hat{c} = \underset{\tilde{c} \in C}{\operatorname{argmin}} \frac{1}{2} \sum_{i=1}^m \left\| w_i \frac{\partial \Pi}{\partial l} (c_0, l_{0i}) (l_i - l_{0i}) \right\|_2^2,$$

$$\text{s.t. } w_i \frac{\partial \Pi}{\partial l} (c_0, l_{0i}) (l_i - l_{0i}) = -w_i \frac{\partial \Pi}{\partial c} (c_0, l_{0i}) (\tilde{c} - c_0) + w_i (im_i - im_{0i}).$$

Этот критерий позволяет оценить жесткое движение камеры \hat{c} , что сводит к минимуму просмотр между эталонной и текущей деформированной фигурой. Эти веса позволяют придавать большее значение некоторым регионам, чем другим, как показано на рис. 3. Следующее выражение используется для вычисления коррекции положения камеры.

На этом рисунке рассматривается контрольная камера светлого цвета, которая видит два опорных треугольника светлого цвета. После деформации вершины движутся в противоположных направлениях на Δl вдоль вертикальной оси. Этот сдвиг вызывает нежесткую деформацию двух треугольников. Здесь рассмотрены два сценария. В левой части рисунка рассмотрены равные веса для обоих треугольников. В этом случае поправка с уравнением (14) не производит никакого движения, так как движение вершин компенсирует друг друга и треугольник имеет одинаковый вес. Правая часть рисунка рассматривает верхний треугольник с весом, в два раза превышающим значение нижнего. В этом случае уравнение (14) возвращает вертикальный сдвиг с половиной Δl . Если $w_2 = 0$, исправление будет представлять собой вертикальный сдвиг вверх в пределах Δl .

Предложение 1. Рассмотрим камеру в некоторой эталонной конфигурации $c_0 \in C$. Камера видит деформированную форму, которая, как предполагается, имеет минимальную энергию растяжения/сжатия. Компенсацию движения жесткой камеры \hat{c} , которая удовлетворяет критерию уравнения (14), можно вычислить следующим образом:

$$(15) \quad \hat{c} = c_0 + W^+ (IM - IM_0), \text{ если } W \text{ - полный ранг столбца.}$$

$$(16) \quad \hat{c} = c_0 + W^T (W W^T)^{-1} (IM - IM_0),$$

в противном случае и если W имеет полный ранг строки.

W^+ является псевдообратной матрицей W .

Доказательство. Принимая во внимание формулу (14) и используя уравнение линеаризации (5), запишем оптимальный критерий как

$$(17) \quad \hat{c} = \operatorname{argmin}_{\tilde{c} \in C} \frac{1}{2} \sum_{i=1}^m \left\| w_i \frac{\partial \Pi}{\partial c}(c_0, l_{0i})(c - c_0) - w_i(im_i - im_{0i}) \right\|_2^2.$$

Это в точности взвешенная сумма квадратов евклидовых норм разности каждого изображения, которое будет отображаться из $c \in C$, и изображения, которое будет отображаться во время выполнения. Эта минимизация может быть переписана в объединенной формуле как

$$(18) \quad \hat{c} = \operatorname{argmin}_{\tilde{c} \in C} \frac{1}{2} \|W(\tilde{c} - c_0) - (IM - IM_0)\|_2^2.$$

Если W является полным рангом столбца, то решение для c этой взвешенной задачи наименьших квадратов будет

$$(19) \quad \hat{c} = c_0 + (W^\top W)^{-1} W^\top (IM - IM_0).$$

Считаем, что в каждый момент времени есть достаточно видимых треугольников, чтобы вычислить коррекцию для шести степеней свободы камеры. Для этого требуем, чтобы были видны по крайней мере два треугольника, чтобы избежать двусмысленных ситуаций [45]. Если задана сетка из N точек, а камера отслеживает полные области этой сетки, справедливо считать, что это условие проверяется в большинстве случаев. В некоторых случаях, когда это может быть не проверено (например, камера слишком близко к объекту), устанавливаем безопасный тест для числа видимых треугольников, который позволяет вычислить решение минимальной нормы. Если W является рангом полной строки, то коррекция камеры вычисляется как минимальная норма движения камеры, удовлетворяющая следующему критерию [47]:

$$(20) \quad \hat{c} = \operatorname{argmin}_{\tilde{c} \in C} \frac{1}{2} \|\tilde{c} - c_0\|_2 \text{ s.t. } W(\tilde{c} - c_0) = (IM - IM_0).$$

Учитывая, что W является рангом полной строки, решение этой задачи с наименьшими квадратами будет

$$(21) \quad \hat{c} = c_0 + W^\top (W W^\top)^{-1} (IM - IM_0).$$

В некоторых изолированных вырожденных случаях, когда W не является ни полным рангом столбца, ни полным рангом строки, камера не корректируется из эталонной конфигурации, и будет $\hat{c} = c_0$.

События, освещенные в этом разделе, можно резюмировать следующим образом. Жесткая компенсация движения нежесткого объекта камерой состоит из решения системы уравнений, которые соотносятся конфигурацией камеры с изображениями треугольников объекта. Если система чрезмерно

ограничена, компенсация состоит из поиска конфигурации, минимизирующей взвешенную сумму остатков. Если нет деформации и треугольники все жесткие, выбор веса не повлияет на результат (игнорируя шум при обнаружении вершин). Если треугольники объекта деформируются, выбор веса повлияет на исправленную конфигурацию \hat{c} , как показано на рис. 3. По этой причине предлагаем использовать эти веса в качестве инструмента для планирования движений на основе треугольной сетки для деформируемых объектов.

4.1. Движение на основе треугольной сетки для визуального обслуживания деформируемых объектов

Определение 1. Давайте рассмотрим камеру, идущую по эталонной траектории и просматривающую деформируемый объект, зацепленный с m треугольниками T_1, \dots, T_m . А треугольная сетка движения состоит из двух основных компонентов:

(i)

$$(22) \quad \gamma : [0, 1] \rightarrow C : s \mapsto \gamma(s),$$

где $[0, 1]$ – нормализованный интервал временного параметра траектории s .

(ii) m – непрерывные положительные вещественные функции w_1, \dots, w_m :

$$(23) \quad w_i : [0, 1] \rightarrow R^+ : s \mapsto w_i(s),$$

$w_i(s) = 0$ для любого s такой, что T_i не видно камерой C , когда камера находится в конфигурации $\gamma(s)$. Непрерывность w_i требуется, чтобы избежать нежелательных прыжков во время коррекции траектории камеры [39].

Указанное выше движение на основе треугольной сетки для компенсации жесткого движения позы камеры заключается в корректировке текущей конфигурации камеры в задаче управления замкнутым контуром путем вычисления поправки из предложения 1 об эталонной конфигурации $\gamma(s)$ и с величинами веса $w_i(s)$ для абсциссы s по траектории. Корректирующая формула может быть написана для данной абсциссы s следующим образом:

1) Если $W(s)$ является полным рангом столбца, то

$$(24) \quad \hat{c}(s) = \gamma(s) + (W(s)^\top W(s))^{-1} W(s)^\top (IM(s) - IM_0(s)).$$

2) Если $W(s)$ не является полным рангом столбца, но является рангом полной строки, то

$$(25) \quad \hat{c}(s) = \gamma(s) + W(s)^\top (W(s) W(s)^\top)^{-1} (IM(s) - IM_0(s)).$$

3) Иначе

$$(26) \quad \hat{c}(s) = \gamma(s).$$

Здесь $W(s)$, $IM(s)$ и $IM_0(s)$ являются расширениями обозначений в уравнениях (11) и (12), когда конфигурация камеры-образца следует по эталонной траектории $c_0(s) = \gamma(s)$, $s \in [0, 1]$. Случай (3) уравнения (26) является граничным и почти никогда не встречается. Он используется в качестве безопасного случая, если когда-либо W не является ни полным рангом столбца, ни полным рангом строки. Вышеприведенные формулы скорректированной траектории $\hat{c}(s)$ из уравнений (24)–(26) можно интерпретировать как удовлетворяющие оптимальному визуальному выравниванию и представляют минимум интеграла по всей траектории разности между эталонным и текущим деформированным видом:

$$(27) \quad J = \int_0^1 \frac{1}{2} \|W(c(s) - \gamma(s)) - (IM(s) - IM_0(s))\|_2^2 ds.$$

Для того, чтобы иметь гладкие поправки необходимо иметь непрерывные веса [39]. Это происходит, например, если веса пропорциональны 2D площади рассматриваемого треугольника на плоскости изображения (в Приложении даны подробные сведения об их использовании и вычислениях). Если объект является органом и задача заключается в компенсации деформаций, наблюдаемых с помощью лапароскопа, то в идеале вес может быть установлен вручную техником перед операцией. Действительно, хирург с помощью техника может указать основную область-мишень органа, которая должна поддерживаться на постоянном уровне при наблюдении лапароскопом. Затем технический специалист может установить более высокие веса для этой области по сравнению с другими. Такой полуавтоматический процесс может быть выполнен при планировании операции.

5. Предлагаемый алгоритм и детали реализации

Алгоритм (1) визуального сервопривода, описанный в этом разделе, использует сетку деформируемого объекта для достижения точного позиционирования камеры относительно объекта, даже когда объект деформируется. Основные шаги следующие:

- 1) Цикл траектории (строки 5–27) выполняет итерацию по опорной траектории камеры, и для каждого положения камеры алгоритм вычисляет:
 - а) для каждого треугольника, используя опорный объект, вес, проекцию опорного изображения и матрицу взаимодействия.
 - б) всеобъемлющую взвешенную матрицу взаимодействия, а также опорные и текущие взвешенные изображения объекта.
- 2) Цикл визуального сервопривода (строки 17–26) затем итеративно находит оптимальное обновление положения камеры, которое минимизирует разницу между опорными и текущими взвешенными изображениями объекта, соответственно обновляя положение камеры.

Algorithm 1 Управление движением на основе треугольной сетки для визуального сервопривода на деформируемом объекте.

Дано: Опорная траектория $\gamma : [0, 1] \rightarrow C : s \mapsto \gamma(s)$.

Дано: Деформируемый объект с сеткой, состоящей из m треугольников T_1, \dots, T_m .

Дано: step size – параметр шага для дискретизации опорной траектории.

Дано: max_{iter} – для максимального числа итераций.

Дано: ϵ_{im} – порог ошибки для остановки на основе разницы изображений.

- 1: **Инициализация:**
 - 2: Загрузить опорную траекторию γ для камеры;
 - 3: Загрузить данные о сетке опорного объекта;
 - 4: Загрузить данные о сетке того же объекта в его деформированном состоянии;
 - 5: **Цикл траектории:**
 - 6: Для $s \in [0, 1]$ **выполнить**
 - 7: Получить текущую конфигурацию камеры $c(s) \leftarrow \gamma(s)$;
 - 8: **Цикл по треугольным сеткам:**
 - 9: По каждому треугольнику i от 1 до m **выполнить**
 - 10: Получить текущий вес w_i , применяя формулы в уравнениях (B14) и (B15);
 - 11: Получить текущую матрицу взаимодействия, применяя формулу из уравнения (7);
 - 12: Получить проекцию опорного изображения im_i^0 треугольного примитива;
 - 13: Вычислить текущую задачу функции треугольника, используя уравнение (10);
 - 14: **Конец цикла**
 - 15: $error \leftarrow \epsilon_{im} + 1$;
 - 16: $iter \leftarrow 0$;
 - 17: **Цикл визуального сервопривода:**
 - 18: Пока выполняется условие $iter < \text{max_iter}$ **or** $error > \epsilon_{im}$ **выполнить**
 - 19: Получить текущую проекцию изображения примитива im_i из текущего положения камеры;
 - 20: Собрать матрицу взаимодействия объекта с учетом веса W с использованием уравнения (11);
 - 21: Собрать опорное и текущее изображения объекта с учетом веса IM, IM_0 с использованием уравнения (12);
 - 22: Решить оптимальную функцию сервопривода, используя уравнения (24)–(26);
 - 23: Использовать полученные $\hat{c}(s)$ для обновления текущей конфигурации камеры;
 - 24: $iter \leftarrow iter + 1$;
 - 25: $error \leftarrow \|W(\hat{c}(s) - c(s)) - (IM - IM_0)\|_2$;
 - 26: **Конец цикла**
 - 27: **Конец цикла**
 - 28: **Вывод:**
 - 29: Окончательная оптимальная траектория визуального сервопривода с учетом текущей деформации объекта.
-

Этот подход позволяет визуальному сервоприводу адаптироваться к деформациям объекта, используя представление сетки для корректировки траектории камеры в соответствии с деформацией.

Этот алгоритм был реализован с использованием Matlab2015a. Он был запущен на ноутбуке с процессором Intel Core (TM) i5-4200U и 6 ГБ оперативной памяти.

Модель объекта, используемая в работе, была треугольной сеткой с 10000 вершинами и 20000 гранями. Опорная траектория была дискретизирована с шагом $step_size$, равным 10 Нз. Максимальное количество итераций, max_iter , было установлено на 100, а порог ошибки для остановки ϵ_{im} был установлен на 0,01 пикселей. Временная сложность составляет $O(N \times m^4 max_iter)$ порядка, а пространственная сложность составляет $O(m^2)$ порядка. Общий алгоритм выполняется со средней скоростью 40 Нз, что позволяет проводить визуальный сервопривод деформируемого объекта в реальном времени.

Этот алгоритм был протестирован в различных сценариях, включая несколько типов объектов, опорных траекторий, режимов деформации и уровней шумов на изображении. В следующем разделе представлены результаты, полученные в этих разнообразных условиях.

6. Результаты по смоделированным данным

Этот раздел демонстрирует несколько примеров проверочных тестов с использованием симулированных данных на двух деформируемых объектах: плоском объекте, который деформируется в форму выпуклости, и модели человеческой печени, которая деформируется на левой и правой долях. Были проведены эксперименты как с линейными, так и с нелинейными деформациями. Метод протестирован на двух геометрических траекториях: линейной и круговой. Для проверки устойчивости предлагаемого подхода был симулирован шум в определении вершин 2D-сетки, с помощью перспективной камеры с фокусным расстоянием 1500. Симуляции выполнялись на процессоре Core (TM) i5-4200U и 6 ГБ оперативной памяти с использованием Matlab2015a. На показанных ниже рисунках расстояния указаны в метрах, скорость перемещения камеры – в метрах в секунду, а скорость вращения – в радианах в секунду. Ошибки в изображении указаны в пикселях.

6.1. Результаты для плоского объекта и линейного геометрического пути

Рисунок 4 представляет целевой объект в трехмерном пространстве и его проекцию на камеру. Объект состоит из 200 треугольников, как показано на рис. 5. При деформации траектория камеры используется в качестве ссылочной зависимости, чтобы гарантировать, что целевые элементы остаются в пределах и динамически корректируются в соответствии с ожидаемым видом. На рис. 6 управляем движением камеры, которое зависит от веса вдоль траектории. Рисунки 5 и 6 представляют соответственно тепловую карту деформации, перемещающуюся от высокой к низкой. В этом эксперименте выделили постоянные веса. Веса являются самыми высокими с непрерывными

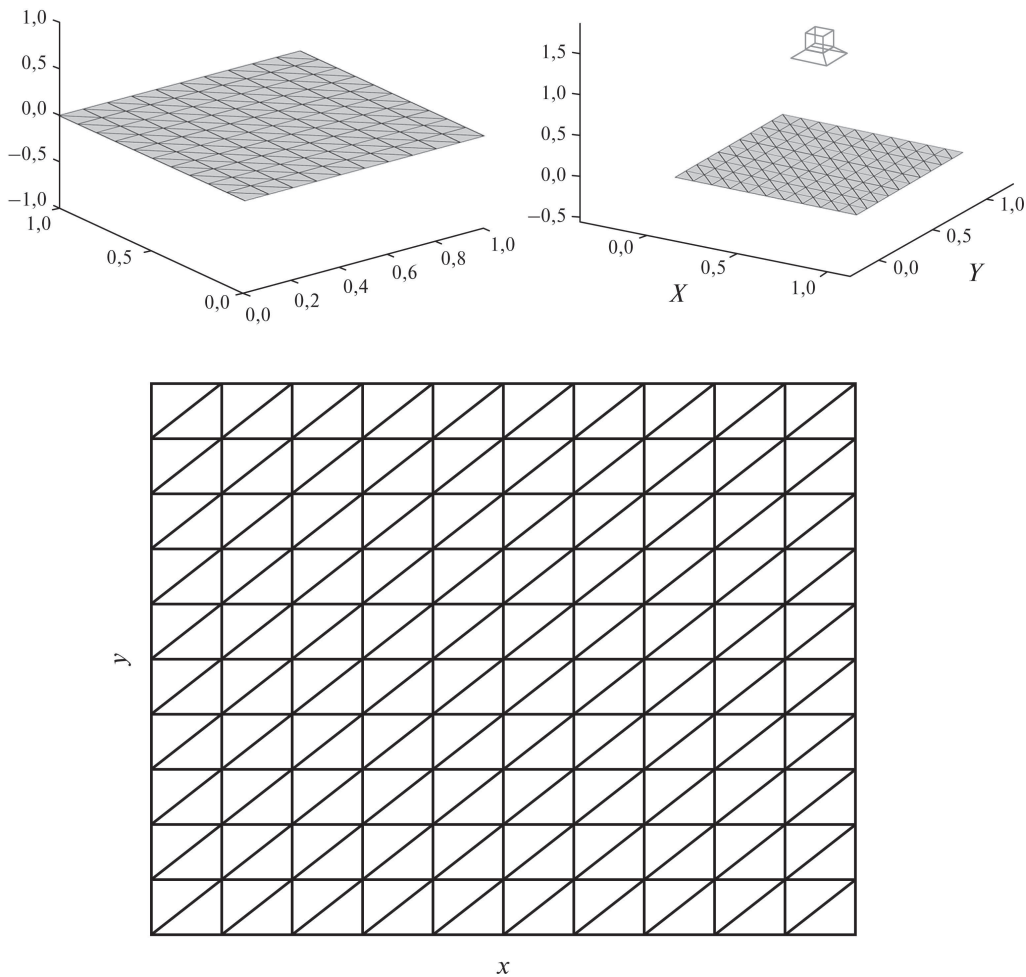


Рис. 4. Справа сверху: опорный объект представляет собой плоскую треугольную 3D-сетку. Слева сверху: эталонный объект и конфигурация камеры в трехмерном пространстве. Внизу: вид плоской сетки в плоскости камеры.

постоянными весами, равными 4. Остальные регионы имеют постоянные веса, равные 1. Не существует верхнего или нижнего предела для выбора веса. Только относительные большие/меньшие значения имеют значение для модуляции большего/меньшего фокуса отслеживания на конкретной области. Области, имеющие значение во время выполнения, имеют большие значения веса для выполнения задачи отслеживания, чтобы лучше следовать за деформированными частями.

6.2. Результаты моделирования печени человека

Печень – самый большой орган в организме человека. Он относится к пищеварительной системе и обеспечивает многие жизненно важные функции организма. В данной работе используется сетка печени для проверки эффек-

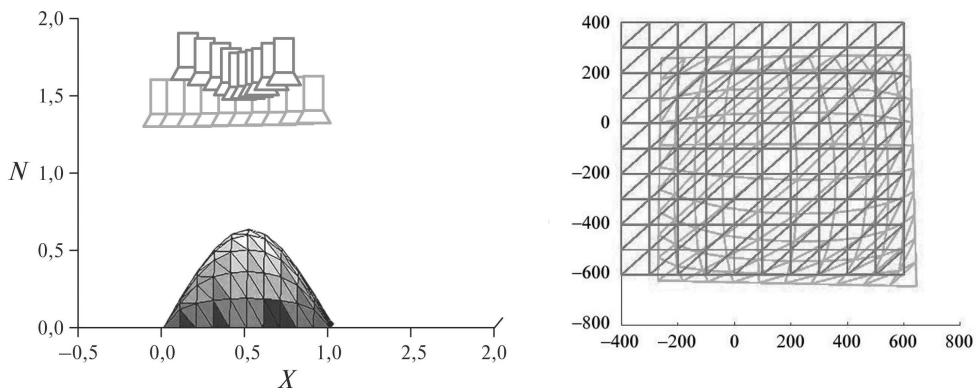


Рис. 5. Слева: деформированный эталонный объект с ориентировочной и деформированной траекториями камеры. Контрольная камера следует линейной запланированной траектории. Светлые камеры – эталонная траектория. Темные камеры – исправленная траектория. Здесь все веса равны единице. Справа: ссылка на просматриваемый объект и деформированный просматриваемый объект. Исправленная траектория должна максимально соответствовать эталонному виду объекта.

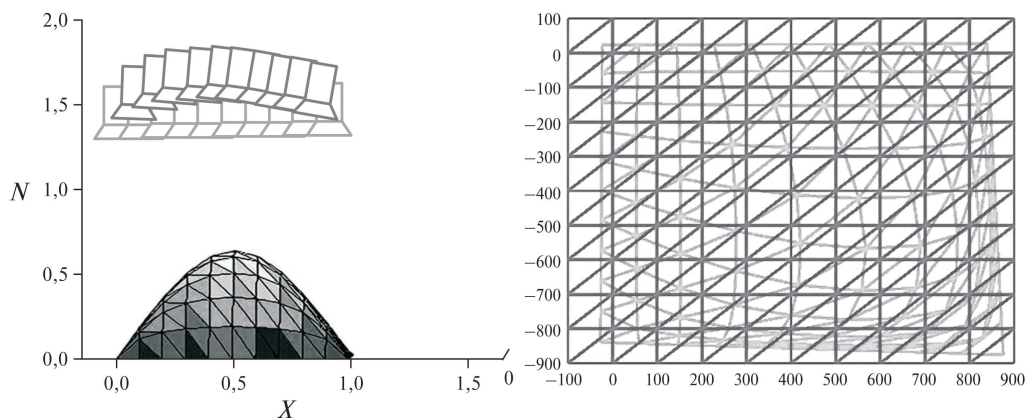


Рис. 6. Слева: деформированный объект 3D-плоскости с камерой, которая следует линейной запланированной траектории. Светлые камеры – эталонная траектория. Темные камеры – исправленная траектория. Здесь все веса равны единице, за исключением весов, связанных с треугольниками на верхнем холме. Они равны 4. Справа: ссылка на просматриваемый объект и деформированный просматриваемый объект. Исправленная траектория должна максимально соответствовать эталонному виду объекта.

тивности используемого метода. Для деформации печени используем гиперэластичную физическую модель Муни–Ривилина и подход, предложенный в [48] для ускорения расчета нелинейных упругих деформаций. Сетка печени содержит 130 382 треугольника и 12 226 вершин (см. рис. 7). Эксперименты проводились по двум типам деформаций: линейной и нелинейной. Для той же сетки печени перемещали выбранную деформированную часть размером 825 на 6 мм по оси z , что считается линейной деформацией. При второй дефор-

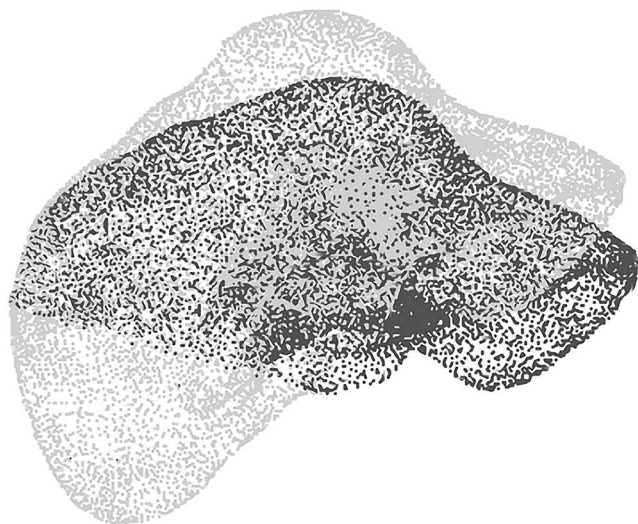
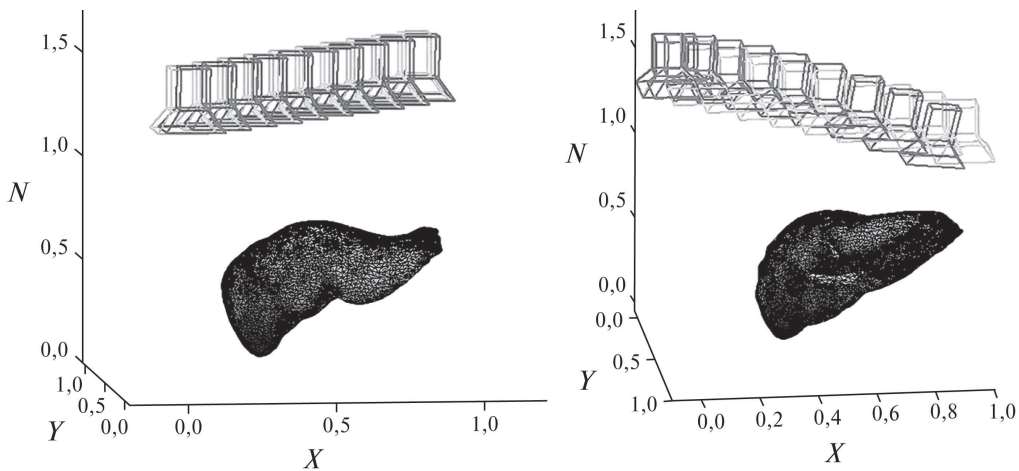


Рис. 7. Темным цветом представлен орган в состоянии покоя. Светлым цветом представлен орган после нелинейной деформации.

мации перемещали те же 825 вершин на 22 мм, что считается нелинейной деформацией.

6.2.1. Результаты с линейной деформацией печени

Для линейной деформации использовали команду для управления движением камеры, которая следовала по ориентировочной горизонтальной прямой линии, как показано на рис. 8. Провели два теста: тест (1), где все веса являются константами и равны 1 для всех треугольников и тест (2), при котором веса изменяются в зависимости от величины видимой деформации. Эта величина вычисляется как отношение площади текущего деформированного треугольника к площади опорного видимого треугольника (см. Приложение). Области вычисляются на видимых треугольниках 2D изображения (опорная и текущая деформированные). Такая стратегия позволяет придавать большее значение деформированным регионам по сравнению с другими областями, менее деформированными или недеформированными. Полученный результат показывает, что камера следила за деформацией печени, не теряя введенной траектории (тест (2)). В тесте (1) траектория камеры изменилась немного и равномерно, так как веса всех треугольников были равны. В тесте (2) камера значительно обновила свою траекторию вокруг деформированной области, сохраняя при этом траекторию, близкую к исходной, далеко от области деформации. Рисунок 9 показывает производительность анализа управления камерой во время динамического отслеживания. Он показывает перемещение и вращение камеры во время отслеживания деформации печени. Скорости вращения ω_y и ω_z одинаковы для обоих тестов. Суммы



Среднее отклонение 2D-точек

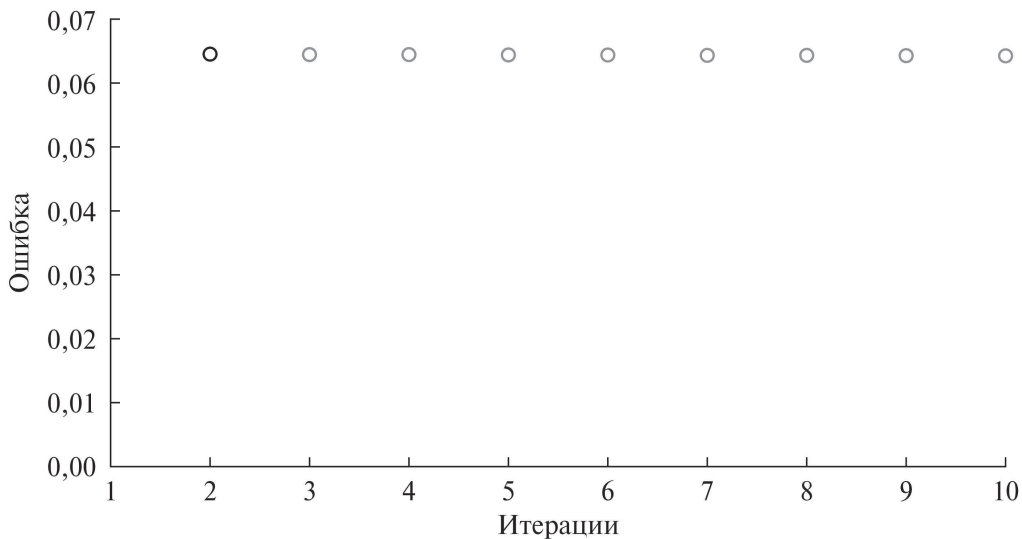


Рис. 8. Печень после линейной упругой деформации и камера следует по ориентировочной прямой траектории. Показаны результаты теста (1). Слева показана эталонная траектория. Справа показана боковая проекция траектории после корректировки. Светлые камеры – эталонная траектория. Темные камеры – исправленная траектория. Внизу показано изменение точек пикселей во время коррекции позы камеры. Темные камеры – эталонная траектория.

перевода по осям y и z также почти одинаковы. Заметная разница заключается в переносе и вращении вдоль оси x , которая соответствует главной оси деформации.

6.2.2. Результаты с нелинейной деформацией печени

В этом эксперименте запустим ту же конфигурацию тестовых (1) и тестовых (2) установок, что и раньше. Рассмотрим нелинейную деформацию с максимальным смещением левой доли печени 6 мм, как показано на рис. 7.

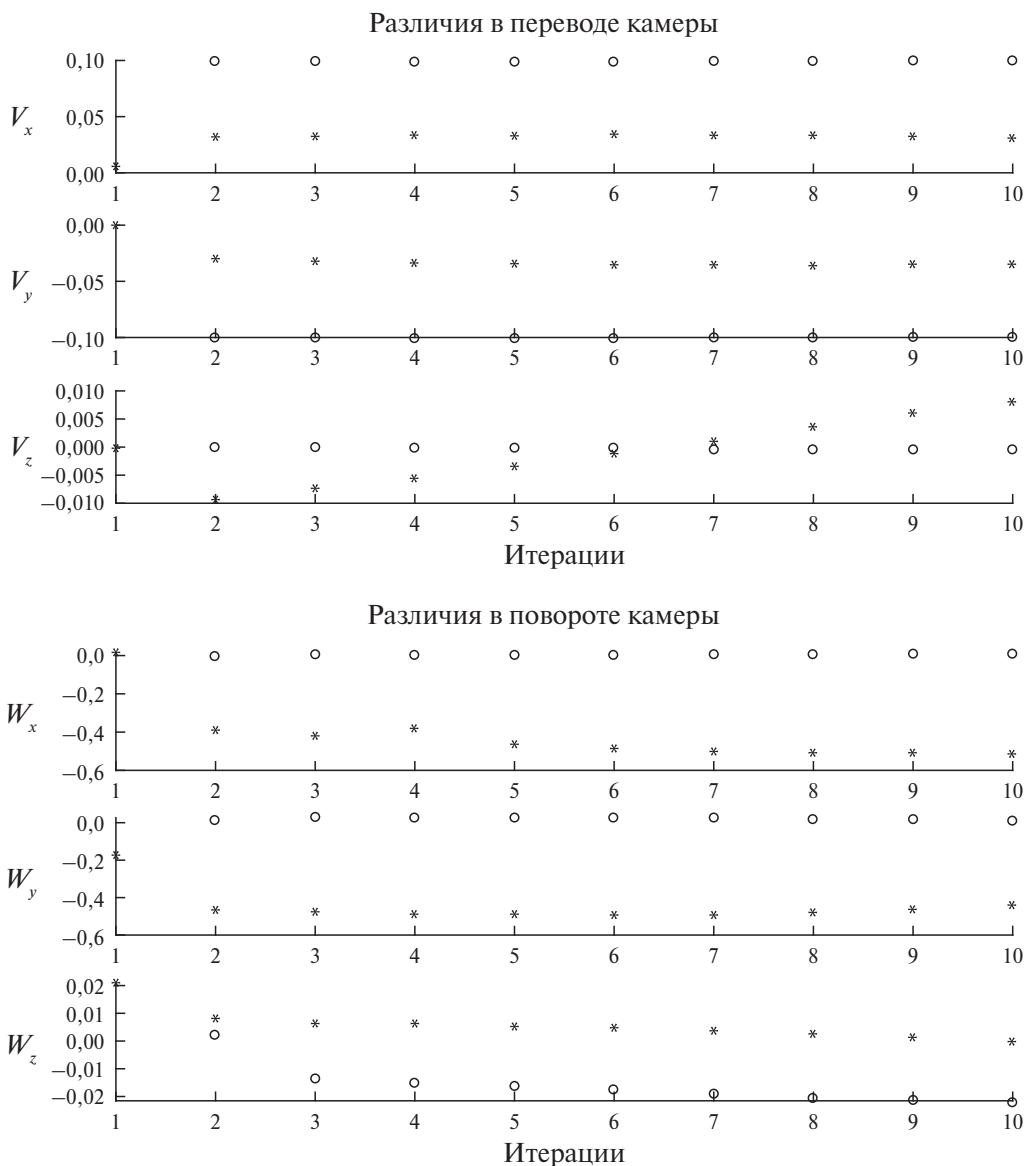


Рис. 9. Анализ производительности скорости камеры при слежении за деформированным объектом с линейной деформацией. Верхняя фигура показывает скорость камеры в переводе. Нижняя фигура показывает скорость камеры при вращении. Круговые точки представляют скорость камеры в соответствии со стратегией тестирования (1). Звездочки обозначают скорость камеры в соответствии со стратегией тестирования (2). ω_y и ω_z схожи для обоих тестов.

Траектория движения камеры представляет собой прямую горизонтальную линию, обозначенную светлым цветом, а скорректированный вид представлен темным цветом. На рис. 10 слева темным цветом показана скорректированная траектория камеры, которая была получена при установке всех весов на 1.

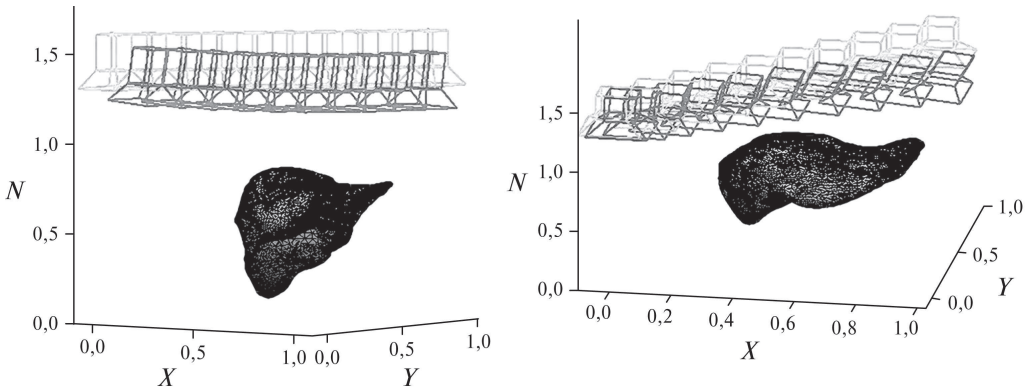


Рис. 10. Печень после нелинейной упругой деформации и камеры следует по ориентировочной прямой траектории. Показаны результаты теста (2). Слева показан вид сбоку. Справа показывает другой вид сбоку. Светлые – эталонная траектория. Темные – исправленная траектория.

На рис. 10 справа темным цветом показана скорректированная траектория движения камеры, полученная с весами, равными 5 для наиболее деформированных областей и равными 1 для наименее деформированных или недеформированных областей. Кривизна скорректированной траектории более важна в случае теста (2) (правый рисунок). Нелинейная деформация треугольников на изображении сохраняется как можно ближе к их областям на эталонном изображении. Весовые функции позволяют брать из печени наиболее релевантные деформированные треугольники и соответствующим образом корректировать объем коррекции. На рис. 10 справа видим, что путь камеры затронут сильнее всего. Это согласуется с нелинейными деформированными треугольниками.

Заметим, что скорость сходимости достаточно высока для подхода (после 10 итераций). Результат расчета коррекции конфигурации камеры показан на рис. 11.

6.2.3. Результаты с нелинейной деформацией и криволинейной опорной траекторией

Для нелинейной траектории выполнили тест на криволинейной плоскости опорной траектории со следующими параметрами: $\theta = (0 : 0.01 : \frac{\pi}{2})$, $c_x = R * \cos(\theta)$, $c_y = R * \cos(\theta)$, $c_z = 1$. Угол камеры остается постоянным во время этой траектории, указывая вниз на отрицательную ось z . Эту конфигурацию эталонной камеры использовали в уравнении (22). Также была протестирована стратегия управления с помощью двух тестов, приведенных в разделе 6.2 (тест (1) и тест (2)). Для теста (1) заметим, что камера следовала по траектории, как показано на рис. 12 справа. Для теста (2) сделали тест с той же стратегией управления и криволинейной траекторией, но изменили вес треугольников, где придали важность для треугольников, которые деформируют постоянное значение веса между 1 и 5, как показано на рис. 6.

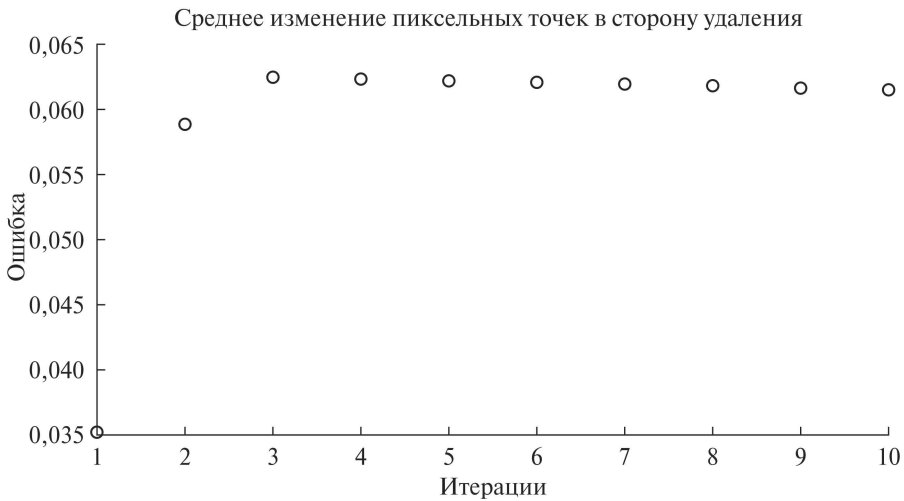
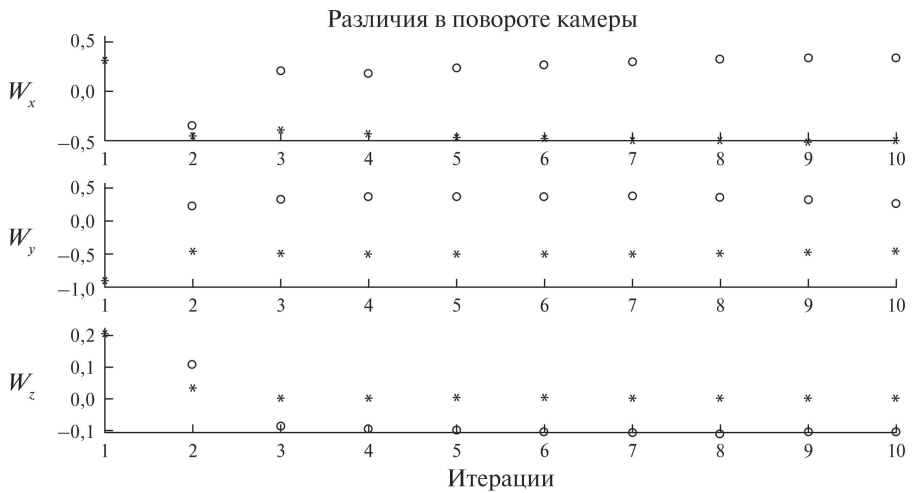


Рис. 11.

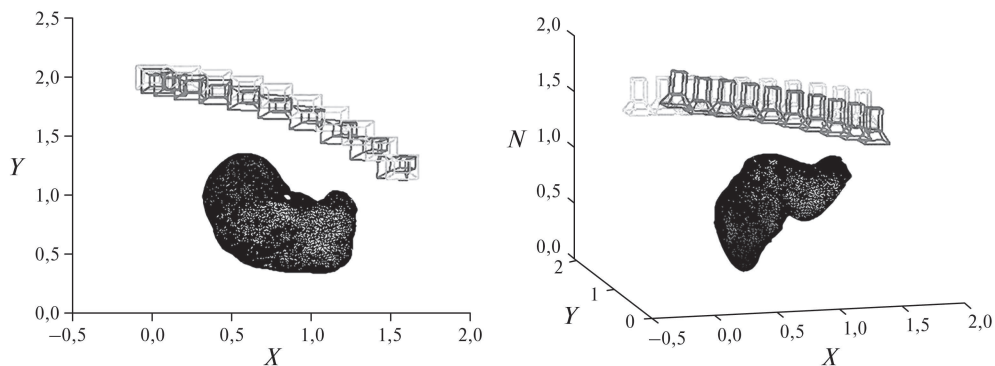


Рис. 12. Печень после нелинейной упругой деформации и камеры следует по опорной криволинейной траектории. Показаны результаты теста (1). Слева показан вид сверху. Справа – вид сбоку. Светлые камеры – эталонная траектория. Темные камеры – исправленная траектория.

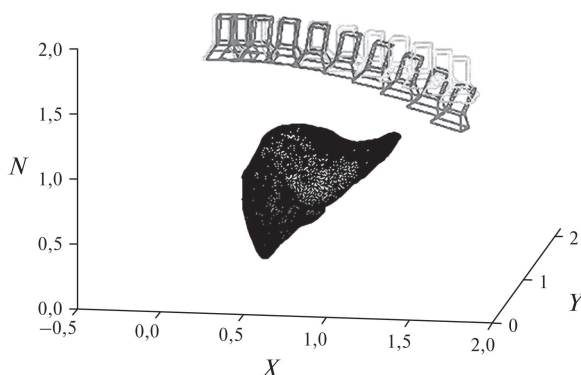


Рис. 13. Печень после нелинейной упругой деформации и камеры следует по опорной криволинейной траектории. Показаны результаты теста (2). Светлые камеры – эталонная траектория. Темные камеры – исправленная траектория.

Полученный результат показывает, что тест (2) лучше теста (1). В тесте (1) камера следовала по траектории, но деформация не была хорошо зафиксирована. Тест 2 успешно прошел по траектории с точным захватом деформации. Кроме того, результат перевода скорости и вращения камеры слежения был лучше в тесте (2), чем в тесте (1), как это показано на рис. 17 (см. ниже). Отметим, что даже в вариации пикселей есть разница, а именно: тест (1) имеет вариацию от 0,06 до 0,1 и тест (2) имеет вариацию от 0 до 0,5.

←

Рис. 11. Анализ производительности скорости камеры при слежении за деформированным объектом с нелинейной деформацией. Верхняя фигура показывает скорость камеры в переводе. Центральная фигура показывает скорость камеры при вращении. Точки представляют скорость камеры в соответствии со стратегией тестирования (1). Звездочки представляют скорость камеры в соответствии со стратегией тестирования (2).

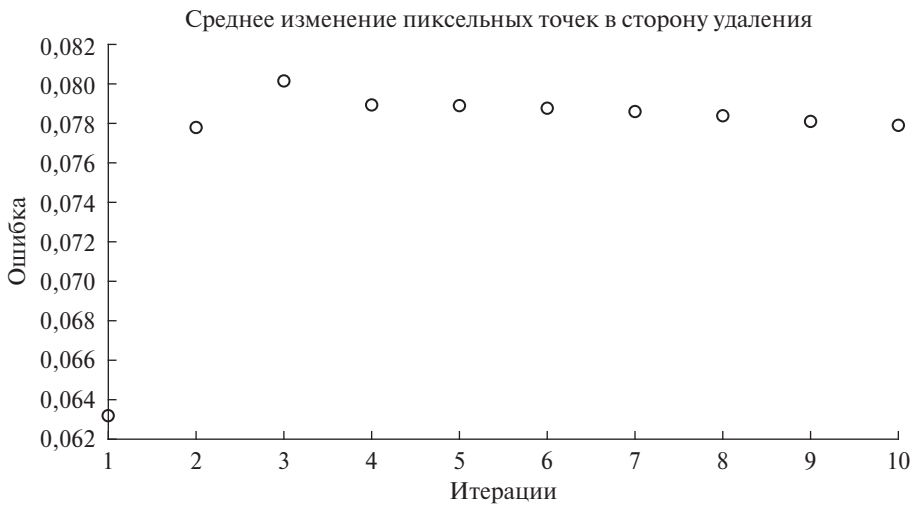
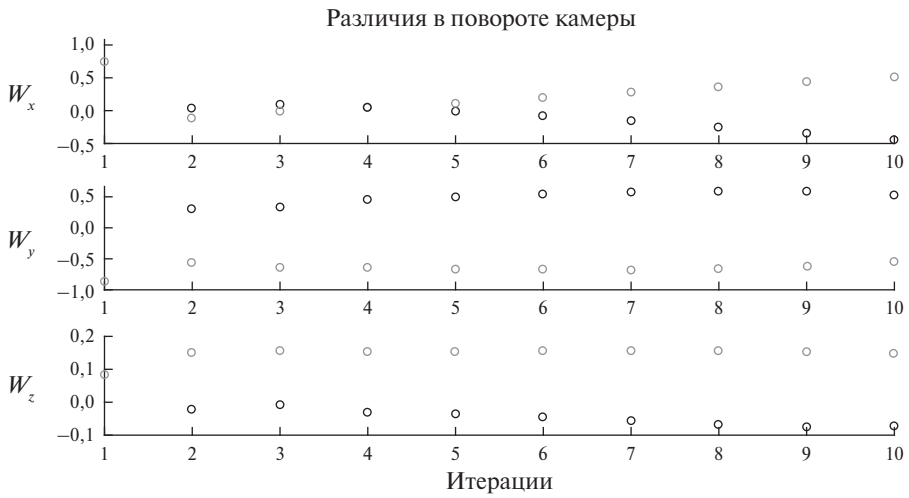
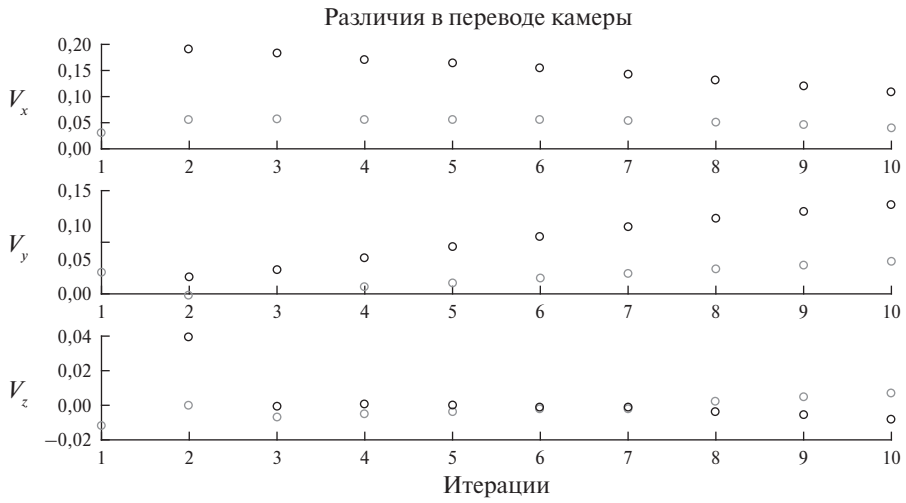


Рис. 14.

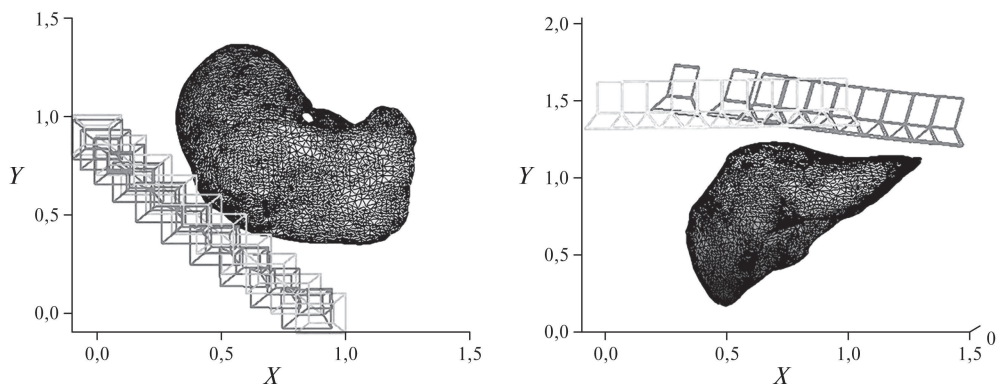


Рис. 15. Печень после нелинейной упругой деформации и камера следует по опорной криволинейной траектории. Показаны результаты теста (1) с шумовыми данными. Слева показан вид сверху. Справа – вид сбоку. Светлые камеры – эталонная траектория. Темные камеры – исправленная траектория.

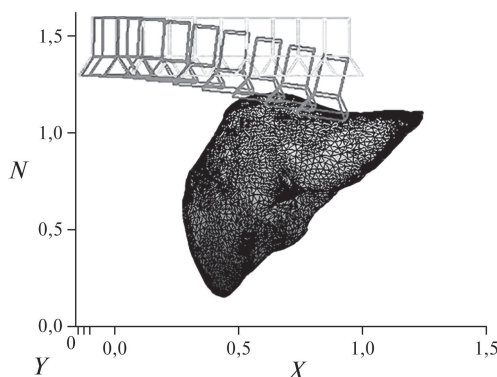


Рис. 16. Печень после нелинейной упругой деформации и камера следует по опорной криволинейной траектории. Показаны результаты теста (2) с шумовыми данными. Светлые камеры – эталонная траектория. Темные камеры – исправленная траектория.

6.3. Устойчивость к шуму при обнаружении 2D вершин

Предлагаемый здесь метод направлен на то, чтобы сделать управление движением камеры более надежным. Поэтому в статье попытались имитировать возмущения измерений реального изображения с помощью гауссового случайного шума, добавленного к уравнению (4). Этот шум центрируется со стандартным отклонением 1 пиксель. Он добавляется к координатам изобра-

Рис. 14. Анализ производительности скорости камеры при слежении за деформированным объектом с нелинейной деформацией и по криволинейной траектории. Верхняя фигура показывает скорость камеры в переводе. Центральная фигура показывает скорость камеры при вращении. Темные точки представляют скорость камеры в соответствии со стратегией тестирования (1). Светлые точки представляют скорость камеры в соответствии со стратегией тестирования (2).

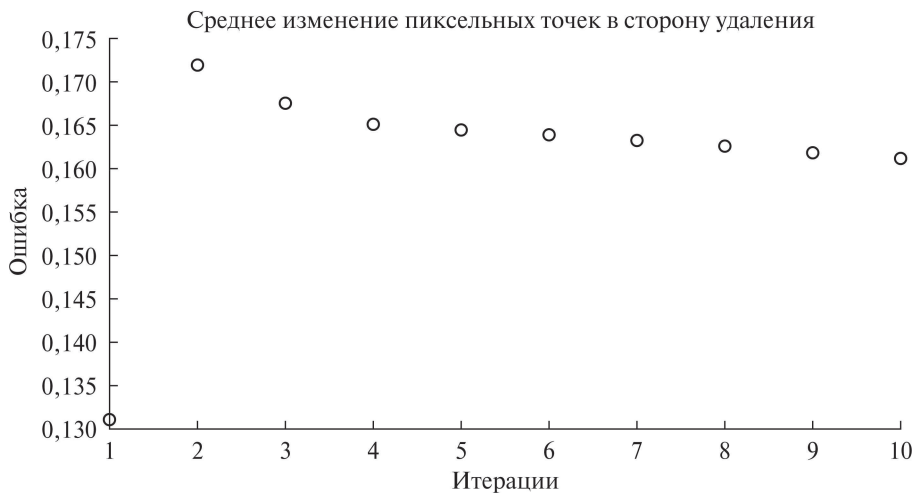
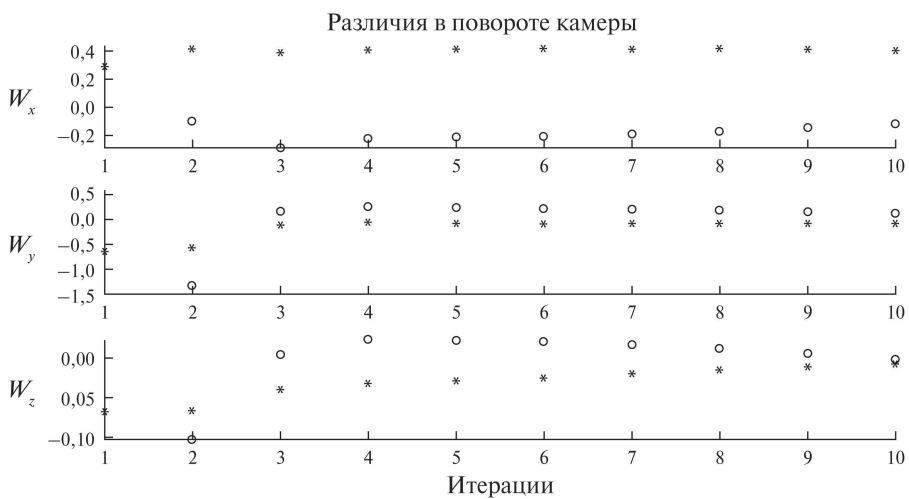


Рис. 17

жения im_i , $1 \leq i \leq m$. Проведен тест на нелинейную деформацию со сравнением экспериментальных установок (как тест (1), так и тест (2)). В результате для теста (1) исправленная траектория была не столь удовлетворительной по сравнению с той же установкой в бесшумовом случае (см. рис. 15). Для теста (2) результат был более совместим с результатом, полученным в бесшумовом случае, как показано на рис. 16. Таким образом, выделение соответствующих весов для соответствующих деформированных регионов не зависит от возмущений при отслеживании и контроле траектории камеры. Рисунок 17 показывает разницу скорости между переносом камеры, поворотом и изменением точки пикселя. Звездочками показаны значения для теста (1), кругами показаны значения для теста (2). Также замечено, что тест (1) переводится и вращается медленнее по сравнению с тестом (2).

6.4. Обсуждения и комментарии

Важность весов четко показана на рис. 5 и 6. С одной стороны, на рис. 6 слева видно, как исходная траектория сильно сгибается, когда она приближается к наиболее деформированной области. Это поведение не наблюдается, когда все веса равны друг другу. С другой стороны, мы видим, как вид деформированной области ближе к эталонному виду на рис. 6 справа по сравнению с рис. 5 справа. Важность структуры и ее надежность можно увидеть в результатах, полученных на рис. 8–17. На этих рисунках подчеркивается надежность предложенного подхода к типу деформации (линейной/нелинейной), типу траекторий (геометрически линейной/нелинейной) и шуму в 2D-обнаружении. Было показано, что в каждом сценарии этих экспериментов предложенный метод позволяет вычислить жесткую коррекцию камеры. Учитывая, что наиболее значимым треугольникам деформируемой области присваиваются большие значения веса, опорная траектория адаптируется в основном к этим областям, независимо от вышеупомянутых категорий.

7. Заключение

В данной статье представлен непараметрический подход к визуальной компенсации нежесткой деформации объекта при жестком движении камеры. Эта компенсация обрабатывается, пока камера движется по опорной траектории благодаря непрерывным весовым функциям. Эти веса позволяют сосредоточиться на наиболее значимых деформирующих областях, одновременно стабилизируя изображение в реальном времени в соответствии с эталонным видом. Разработанная в статье стратегия компенсации встраивает

Рис. 17. Анализ производительности скорости камеры при слежении за деформированным объектом с нелинейной деформацией и шумовыми данными. Верхняя часть рисунка показывает скорость камеры в переводе. Центральная часть рисунка показывает скорость камеры при вращении. Точки обозначают скорость камеры в соответствии со стратегией теста (1). Звездочками обозначена скорость камеры в соответствии со стратегией теста (2).

ся в оптимальную систему визуального контроля. Экспериментально доказано, что такой подход устойчив к линейным и нелинейным деформациям, линейным и нелинейным геометрическим траекториям, а также к шуму в 2D-обнаружении. Подчеркиваем, что предлагаемый метод не требует информации о типе объекта, деформации или траектории.

Эта работа позволяет создать прочную формальную основу для перспективного реального применения в контексте лапароскопии. Хирург может спланировать вмешательство с помощью технического специалиста, придавая значительный вес важным областям целевого органа и описывая эталонную геометрическую траекторию для лапароскопа. Во время операции обеспечение движения может быть выполнено роботом, держащим лапароскоп, с использованием изложенного метода. На основе современных подходов к обнаружению 2D-ориентиров в лапароскопии, алгоритм реализуется в реальном времени для обнаружения и построения 2D-сетки с вершинами. Это позволит полностью автоматизировать отслеживание обработанной траектории и визуальное сервоприводное управление.

ПРИЛОЖЕНИЕ

Вычисление матрицы взаимодействия

Если дана 3D-вершина с координатами $(X, Y, Z)^T$ в кадре камеры, то ее 2D-проекция на плоскость камеры задается как

$$(П.1) \quad x = X/Z,$$

$$(П.2) \quad y = Y/Z,$$

где $im = (x, y)^T$ – координата изображения трехмерной точки без учета фокусного расстояния и главной точки (в данной работе рассматриваем калиброванную камеру с известными внутренностями). Производная по времени приведенной выше проекции относительно опорной 3D координаты вершины (X_0, Y_0, Z_0) будет

$$(П.3) \quad \dot{x} = \dot{X}/Z - \dot{Z}x/Z,$$

$$(П.4) \quad \dot{y} = \dot{Y}/Z - \dot{Z}y/Z.$$

Запишем скорость перемещения камеры как $v = (v_x, v_y, v_z)^T$, тогда $\omega = (\omega_x, \omega_y, \omega_z)^T$ таким образом, чтобы полный вектор скорости камеры записывался как $\dot{c} = (v^T, \omega^T)^T$. Формула, связывающая скорость 3D-точки со скоростями камеры, может быть записана по следующей формуле из классической работы [49]:

$$(П.5) \quad \dot{X} = -v_x - \omega_y Z + \omega_z Y,$$

$$(П.6) \quad \dot{Y} = -v_y - \omega_z X + \omega_x Z,$$

$$(П.7) \quad \dot{Z} = -v_z - \omega_x Y + \omega_y X.$$

Заменяем выражения трехмерной скорости уравнений (П.5)–(П.7) выражениями двумерной скорости (П.3)–(П.4)

$$(П.8) \quad \dot{x} = -v_x/Z + xv_z/Z + xy\omega_x - (1 + x^2)\omega_y + y\omega_z,$$

$$(П.9) \quad \dot{y} = -v_y/Z + yv_z/Z + (1 + y^2)\omega_x - xy\omega_y - x\omega_z.$$

Переупорядочив термины и используя обозначение матрицы взаимодействия из уравнения (6), можно переписать указанную выше систему как

$$(П.10) \quad (\dot{x}, \dot{y})^\top = L(x, y, Z)\dot{c}.$$

Если рассмотрим достаточно высокий период выборки T (по крайней мере 30 кадров в секунду), то вышеприведенную формулу можно дискретизировать следующим образом:

$$(П.11) \quad (x - x_0, y - y_0)^\top / T = L(x, y, Z)(c - c_0) / T,$$

где c представлен позицией перевода центра камеры и вектором поворота Родригеса

$$(П.12) \quad c = (t_x, t_y, t_z, \theta_x, \theta_y, \theta_z)^\top.$$

Здесь (t_x, t_y, t_z) – координаты трехмерной позиции, $\theta = \sqrt{\theta_x^2 + \theta_y^2 + \theta_z^2}$ – угол поворота и $(\theta_x/\theta, \theta_y/\theta, \theta_z/\theta)$ единица измерения оси вращения [50]. $c_0 = (t_x^0, t_y^0, t_z^0, \theta_x^0, \theta_y^0, \theta_z^0)^\top$ является конфигурацией эталонной камеры, близкой к c . Упрощающее уравнение (П.11) по периоду выборки с обеих сторон дает

$$(П.13) \quad (x - x_0, y - y_0)^\top = L(x, y, Z)(c - c_0).$$

Это отображение известно как матрица взаимодействия, которая связывает близкие вариации скорости камеры с близкими вариациями 2D проецируемых точек. Рассмотрим отображение Π , которое определено в уравнении (3), предполагающем проекцию 3D треугольника как составного геометрического примитива 3 вершин. В уравнении (5) представлено разложение Тейлора первого порядка по формуле (4) о конфигурации эталонной камеры c_0 , близкой к заданной конфигурации камеры c . Термин $\frac{\partial \Pi}{\partial c}$ – это матрица взаимодействия, которая связывает вариацию конфигурации камеры с 2D-проекцией 3 вершин, составляющих треугольник. Таким образом, аналитическое выражение $\frac{\partial \Pi}{\partial c}$ получается путем вертикального сцепления матрицы, приведенной в уравнении (П.10). Таким образом, матрица $\frac{\partial \Pi}{\partial c}$ состоит из 6 столбцов. В ней 6 строк, так как каждая вершина треугольника дает 2 уравнения, а в каждом треугольнике 3 вершины.

Пример вычисления веса с видимыми областями треугольников

Не существует оптимального или точного способа вычисления весовых функций [39]. Тем не менее существует множество подходов, которые могут быть использованы для построения функций непрерывного веса. Самое

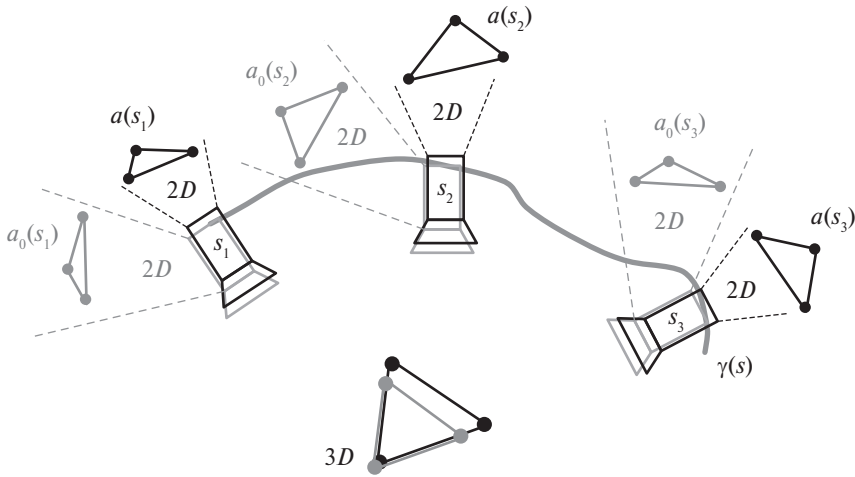


Рис. 18. Пример вычисления веса на основе площади двумерных треугольников.

простое – сделать их постоянными, как это было сделано в протокольном тесте (1) в описанном эксперименте. Другой метод основан на зависимости от степени деформации, что позволяет уделить деформированным участкам большее внимание, чем остальным. В этом случае предположим, что есть один недеформированный 3D треугольник, рассматриваемый камерой по опорной траектории $\gamma(s)$. Двухмерная проекция этого треугольника на камеру-образец непрерывна вдоль $\gamma(s)$, поскольку это статический объект, который видит непрерывно движущаяся камера (см. рис. 18). На этом рисунке рассматривается эталонная камера в светлом цвете с одним эталонным треугольником в светлом цвете вдоль эталонной траектории $\gamma(s)$, $0 \leq s \leq 1$. После деформации новый деформированный треугольник показан в 3D пространстве темным цветом. Исправленная камера по траектории отображается темным цветом. Были взяты три образца траектории s_1 , s_2 и s_3 , чтобы изобразить три пары видов как эталонного треугольника от камеры наблюдения, так и деформированного треугольника от скорректированной позы камеры. Как видно из плоскостей 2D-изображений обеих камер, площадь рассматриваемых треугольников непрерывно изменяется вдоль траектории. Здесь вес вычисляется как $w(s) = a_0(s) \text{abs}(a_0(s) - a(s)) / (1 + a_0(s))$. Двухмерное расположение спроецированного треугольника может быть записано как

$$(II.14) \quad im_0(s) = (x_1^0(s), x_2^0(s), x_3^0(s), y_1^0(s), y_2^0(s), y_3^0(s))^T.$$

Площадь спроецированного треугольника вдоль $\gamma(s)$ также непрерывна и может быть вычислена как каждый s как

$$(II.15) \quad a_0(s) = \frac{1}{2} \det \begin{pmatrix} x_1^0(s) & y_1^0(s) & 1 \\ x_2^0(s) & y_2^0(s) & 1 \\ x_3^0(s) & y_3^0(s) & 1 \end{pmatrix},$$

где $\det()$ обозначает определитель квадратичных матриц. Если рассмотрим $a(s)$ как площадь проецируемого треугольника на камеру во время выполнения, то формула вычисления весовой функции может быть вычислена как

$$(П.16) \quad w(s) = a_0(s) \frac{\text{abs}(a_0(s) - a(s))}{(1 + a_0(s))},$$

где $\text{abs}()$ обозначает абсолютное значение вещественных чисел. Вышеуказанный $w(s)$ является непрерывным и состоит из непрерывных функций ($a(s)$ и $a_0(s)$). Он положителен на каждые s для $0 \leq s \leq 1$. Если треугольник не виден в ссылочном представлении, его область $a_0(s)$ равна NULL, что делает его вес также нулевым. Чем больше треугольник деформирован, тем больше $\text{abs}(a_0(s) - a(s))$, что позволяет учитывать больше деформированных областей.

СПИСОК ЛИТЕРАТУРЫ

1. *Petit A., Lippiello V., Fontanelli G.A., Siciliano B.* Tracking elastic deformable objects with an RGB-d sensor for a pizza chef robot // *Robotics and Autonomous Systems*. 2017. V. 88. P. 187–201. <https://doi.org/10.1016/j.robot.2016.08.023>. Accessed 2021-06-17.
2. *Haouchine N., Dequidt J., Peterlik I., Kerrien E., Berger M.-O., Cotin S.* Image-guided simulation of heterogeneous tissue deformation for augmented reality during hepatic surgery // *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, P. 199–208. <https://doi.org/10.1109/ISMAR.2013.6671780>
3. *Rastegarpanah A., Aflakian A., Stolkin R.* Optimized hybrid decoupled visual servoing with supervised learning / *Proceedings of the Institution of Mechanical Engineers, Part I // J. Syst. Control Engin.* 2022. V. 236(2). P. 338–354. <https://doi.org/10.1177/09596518211028379>. Publisher: IMECHE. Accessed 2021-07-21.
4. *Chi C., Berenson D.* Occlusion-robust deformable object tracking without physics simulation // *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2019. P. 6443–6450. <https://arxiv.org/abs/2101.007332101.00733>. <https://doi.org/10.1109/IROS40897.2019.8967827>. Accessed 2021-06-29.
5. *Lagneau R.* Shape Control of Deformable Objects by Adaptive Visual Servoing // *INSA de Rennes*, 2020. <https://tel.archives-ouvertes.fr/tel-03087518> Accessed 2021-06-17.
6. *Lagneau R., Krupa A., Marchal M.* Active Deformation through Visual Servoing of Soft Objects // *2020 IEEE International Conference on Robotics and Automation (ICRA)*, P. 8978–8984 (2020). <https://doi.org/10.1109/ICRA40945.2020.9197506>. ISSN: 2577-087X
7. *Chaumette F. Ikeuchi K.* (ed.) *Visual Servoing*. Springer. Boston. 2014. P. 869–874.
8. *Tahri O., Chaumette F.* Image moments: generic descriptors for decoupled image-based visual servo // *IEEE International Conference on Robotics and Automation*, 2004. *Proceedings. ICRA '04*. 2004. V. 2. P. 1185–11902. <https://doi.org/10.1109/ROBOT.2004.1307985> ISSN: 1050-4729

9. *Agravante D.J., Claudio G., Spindler F., Chaumette F.* Visual servoing in an optimization framework for the whole-body control of humanoid robots // *IEEE Robotics and Automation Letters*. V. 2. No. 2. P. 608–615.
<https://doi.org/10.1109/LRA.2016.2645512>
10. *Ren X., Li H., Li Y.* Image-based visual servoing control of robot manipulators using hybrid algorithm with feature constraints // *IEEE Access* 8. 2020. P. 223495–223508. <https://doi.org/10.1109/ACCESS.2020.3042207>
11. *Wang T., Wang W., Wei F.* An overview of control strategy and trajectory planning of visual servoing / Fei, M., Li, K., Yang, Z., Niu, Q., Li, X. (eds.) *Recent Featured Applications of Artificial Intelligence Methods // LSMS 2020 and ICSEE 2020 Workshops*. P. 358–370. Springer, Singapore 2020
12. *Corke P., Hutchinson S.* A new partitioned approach to image-based visual servo control // *IEEE Transactions on Robotics and Automation*. V. 17. No. 4. P. 507–515.
<https://doi.org/10.1109/70.954764>
13. *Nicholas R. Gans, Seth A.* Performance Tests for Visual Servo Control Systems, with Application to Partitioned Approaches to Visual Servo Control Hutchinson, Peter I. Corke // *Int. J. Robot. Res.* 2003.
<https://journals.sagepub.com/doi/abs/10.1177/027836490302210011> Accessed 2021-03-07
14. *Janabi-Sharifi F., Wilson W.J.* Automatic selection of image features for visual servoing // *IEEE Transactions on Robotics and Automation*. 1997. V. 13. No. 6. P. 890–903. <https://doi.org/10.1109/70.650168>
15. *Chaumette F.* Image moments: a general and useful set of features for visual servoing // *IEEE Transactions on Robotics*. 2024. V. 20. No. 4. P. 713–723.
<https://doi.org/10.1109/TRO.2004.829463>
16. *Molnar C., Nagy T.D., Elek R.N., Haidegger T.* Visual servoing-based camera control for the da vinci surgical system // *2020 IEEE 18th International Symposium on Intelligent Systems and Informatics (SISY)*. 2020. P. 107–112.
<https://doi.org/10.1109/SISY50555.2020.9217086>
17. *Mohamed I.* MPPI-VS: Sampling-Based Model Predictive Control Strategy for Constrained Image-Based and Position-Based Visual Servoing // *IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2021.
<https://doi.org/10.48550/arXiv.2104.04925>
18. *Lagneau R., Krupa A., Marchal M.* Active deformation through visual servoing of soft objects // *2020 IEEE International Conference on Robotics and Automation (ICRA)*. 2020. P. 8978–8984. <https://doi.org/10.1109/ICRA40945.2020.9197506>
19. *Hu Z., Han T., Sun P., Pan J., Manocha D.* 3-d deformable object manipulation using deep neural networks // *IEEE Robotics and Automation Letters*. 2019. V. 4. No. 4. P. 4255–4261. <https://doi.org/10.1109/LRA.2019.2930476>
20. *Jia B., Hu Z., Pan J., Manocha D.* Manipulating highly deformable materials using a visual feedback dictionary // *2018 IEEE International Conference on Robotics and Automation (ICRA)*. 2018. P. 239–246.
<https://doi.org/10.1109/ICRA.2018.8461264>

21. *Hu Z., Sun P., Pan J.* Three-dimensional deformable object manipulation using fast online gaussian process regression // *IEEE Robotics and Automation Letters*. 2018. V. 3. No. 2. P. 979–986. <https://doi.org/10.1109/LRA.2018.2793339>
22. *Zhu J.* Vision-based robotic manipulation of deformable linear objects. PhD thesis, Université Montpellier. 2020.
23. *Chen Z., Li S., Zhang N., Hao Y., Zhang X.* Eye-to-hand robotic visual tracking based on template matching on fpgas // *IEEE Access*. 2019. V. 7. P. 88870–88880. <https://doi.org/10.1109/ACCESS.2019.2926807>
24. *Staneva V., Younes L.* Modeling and estimation of shape deformation for topology-preserving object tracking // *SIAM J. Imag. Sci.*. 2014. V. 7. No. 1. P. 427–455. <https://doi.org/10.1137/130919714>
25. *Hu Y., Carter T.J., Ahmed H.U., Emberton M., Allen C., Hawkes D.J., Barratt D.C.* Modelling prostate motion for data fusion during image-guided interventions // *IEEE Transactions on Medical Imaging*. V. 30. No. 11. P. 1887–1900. <https://doi.org/10.1109/TMI.2011.2158235>
26. *Chen Q., Sun Q.-S., Heng P.A., Xia D.-S.* Two-stage object tracking method based on kernel and active contour // *IEEE Transactions on Circuits and Systems for Video Technology*. 2020. V. 20. No. 4. P. 605–609. <https://doi.org/10.1109/TCSVT.2010.2041819>
27. *Cao X., Lan J., Rong Li X.* Extension-deformation approach to extended object tracking // 2016 19th International Conference on Information Fusion. (FUSION). 2016. P. 1185–1192.
28. *Joo H., Simon T., Sheikh Y.* Total capture: A 3d deformation model for tracking faces, hands, and bodies // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2018.
29. *Royer L., Marchal M., Le Bras A., Dardenne G., Krupa A.* Real-time tracking of deformable target in 3d ultrasound images // 2015 IEEE International Conference on Robotics and Automation (ICRA). P. 2430–2435. <https://doi.org/10.1109/ICRA.2015.7139523>. ISSN: 1050-4729.
30. *Royer L., Krupa A., Dardenne G., Bras A.L., Marchand E., Marchal M.* Real-time target tracking of soft tissues in 3d ultrasound images based on robust visual information and mechanical simulation // *Medical Image Analysis*. 2017. V. 35. P. 582–598. ISSN 1361-8415. <https://doi.org/10.1016/j.media.2016.09.004>
31. *Kajihara K., Huang S., Bergstrom N., Yamakawa Y., Ishikawa M.* Tracking of trajectory with dynamic deformation based on dynamic compensation concept // 2017 IEEE International Conference on Robotics and Biomimetics (ROBIO). 2017. P. 1979–1984. <https://doi.org/10.1109/ROBIO.2017.8324709>
32. *Zhou H., Ma J., Tan C.C., Zhang Y., Ling H.* Cross-weather image alignment via latent generative model with intensity consistency // *IEEE Transactions on Image Processing*. 2020. V. 29. P. 5216–5228. <https://doi.org/10.1109/TIP.2020.2980210>
33. *Toriya H., Dewan A., Kitahara I.* Sar2opt: Image alignment between multi-modal images using generative adversarial networks // *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*. 2019. P. 923–926. <https://doi.org/10.1109/IGARSS.2019.8898605>

34. *Shingo Kagami, K.H. Kotaro Omi*: Alignment of a flexible sheet object with position-based and image-based visual servoing // *Advanced Robotics*. 2016. V. 30. P. 965–978. <https://doi.org/10.1080/01691864.2016.1183518>
35. *Xi Shen, A.A.E.M.A. Francois Darmon*. Ransac-flow: Generic two-stage image alignment // *Computer Vision-ECCV 2020*. V. 12349. P. 618–637.
36. *Dong Y., Liang T., Zhang Y., Du B.* Spectral spatial weighted kernel manifold embedded distribution alignment for remote sensing image classification // *IEEE Transactions on Cybernetics*. (2021). V. 51(6). P. 3185–3197. <https://doi.org/10.1109/TCYB.2020.3004263>
37. *Mathiassen K., Glette K., Elle O.J.* Visual servoing of a medical ultrasound probe for needle insertion // *2016 IEEE International Conference on Robotics and Automation (ICRA)*. P. 3426–3433. <https://doi.org/10.1109/ICRA.2016.7487520>
38. *Mura M., Abu-Kheil Y., Ciuti G., Visentini-Scarzanella M., Menciassi A., Dario P., Dias J., Seneviratne L.* Vision-based haptic feedback for capsule endoscopy navigation: a proof of concept // *Micro-Bio Robot*. V. 11. P. 35–45. <https://doi.org/10.1007/s12213-016-0090-2> Accessed 2021-03-23.
39. *Malti A., Taïx M., Lamiraux F.* A general framework for planning landmark-based motions for mobile robots // *Advanced Robotics*. 2011. V. 25. No. (11–12). P. 1427–1450. <https://doi.org/10.1163/016918611X579457>
40. *Sengupta A., Krupa A., Marchand E.* Visual Tracking of Deforming Objects Using Physics-based Models // *ICRA 2021 – IEEE International Conference on Robotics and Automation 2021*.
41. *Feng X., Mei W., Hu D.* A Review of Visual Tracking with Deep Learning // *Atlantis Press*. (2016). P. 231–234. ISSN: 1951-6851.
42. *Marvasti-Zadeh S.M., Cheng L., Ghanei-Yakhdan H., Kasaei S.* Deep Learning for Visual Tracking: A Comprehensive Survey // *IEEE Transact. Intelligent Transport*. V. 23. No. 5. 2021. P. 1–26.
43. *Malti A., Hartley R., Bartoli A., Kim J.-H.* Monocular template-based 3d reconstruction of extensible surfaces with local linear elasticity // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2013.
44. *Malti A., Bartoli A., Hartley R.* A linear least-squares solution to elastic shape-from-template // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2015.
45. *Malti A., Herzet C.* Elastic shape-from-template with spatially sparse deforming forces // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017.
46. *Casillas-Perez D., Pizarro D., Fuentes-Jimenez D., Mazo M., Bartoli A.* Equiareal Shape-from-Template // *J. Math. Imaging Vision*. 2019. V. 61. No. 5. P. 607–626.
47. *Ben-Israel A., Greville T.N.E.* *Generalized Inverses Theory and Applications* // Springer. 2003. Google-Books-ID: o_zXUXaqGU8C
48. *Saidi F., Malti A.* Fast and accurate nonlinear hyper-elastic deformation with a posteriori numerical verification of the convergence of solution: Application to the simulation of liver deformation // *Numerical Methods in Biomedical Engineering*. 2021. V. 37. No. 5. <https://doi.org/10.1002/cnm.3444>

49. *Chaumette F., Hutchinson S.* Visual servo control. I. basic approaches // IEEE Robot. Autom. Magaz. 2006 V. 13. No. 4. P. 82–90.
<https://doi.org/10.1109/MRA.2006.250573>
50. *Murray R.M., Li Z., Sastry S.S., Sastry S.S.* A Mathematical Introduction to Robotic Manipulation // CRC Press. 1994. <https://doi.org/10.1201/9781315136370>

Статья представлена к публикации членом редколлегии Д.В. Виноградовым.

Поступила в редакцию 13.02.2024

После доработки 26.08.2024

Принята к публикации 26.09.2024