МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ  
РОССИЙСКОЙ ФЕДЕРАЦИИ

федеральное государственное автономное   
образовательное учреждение высшего образования  
«Самарский национальный исследовательский университет   
имени академика С.П. Королева»  
(Самарский университет)

Институт информатики и кибернетики

Кафедра технической кибернетики

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

Применение алгоритмов машинного обучения для задач захвата движения человека на видеоизображении

по программе бакалавриата по направлению подготовки   
01.03.02 Прикладная математика и информатика,

профиль «Компьютерные науки»

Обучающийся А.А. Сорока

(*подпись*)

Научный руководитель ВКР,

*доцент, к.ф.-м.н.*

Д.А. Савельев

(*подпись*)

Нормоконтролёр С.В. Суханов

(*подпись*)

Самара 2024

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ  
РОССИЙСКОЙ ФЕДЕРАЦИИ

федеральное государственное автономное   
образовательное учреждение высшего образования  
«Самарский национальный исследовательский университет   
имени академика С.П. Королева»

Институт информатики и кибернетики

Кафедра технической кибернетики

|  |  |
| --- | --- |
|  | УТВЕРЖДАЮ  Заведующий кафедрой  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ А.В. Куприянов  «\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 20\_\_\_\_ г. |
|  |  |

**ЗАДАНИЕ НА ВЫПУСКНУЮ КВАЛИФИКАЦИОННУЮ РАБОТУ   
БАКАЛАВРА**

обучающемуся группы 6409-010302D ***Сорока Александру Александровичу***

Тема ВКР: ***Применение алгоритмов машинного обучения для задач захвата движения человека на видеоизображении***

утверждена приказом по университету от « 08 » 04 2021 г. № 333-ст .

Исходные данные: *двумерное волновое уравнение, метод разделения переменных Фурье, признаки сходимости функциональных рядов, метод вычислительного эксперимента, метод конечных разностей*.

Перечень вопросов, подлежащих разработке в ВКР:

* *получение решения краевой задачи для волнового уравнения в виде бесконечного ряда Фурье при различных способах описания входного импульса;*
* *получение оценки остатка ряда Фурье;*
* *разработка программного средства для численного моделирования и исследования погрешности;*
* *обеспечение контроля погрешности решения волнового уравнения;*
* *разработка программы разностного решения волнового уравнения и использование программного средства численного моделирования с помощью рядов Фурье для тестирования программы разностного решения.*

|  |  |
| --- | --- |
| Руководитель ВКР  *доцент*  Д.А. Савельев  (*подпись*)  « 15 » 02 2022 г. | Задание принял к исполнению  А.А. Сорока  (*подпись*)  « 15 » 02 2022 г. |

*Примечание: Задание в данном случае печатается* ***с двух сторон одного листа****, пункты задания и все подписи будут находиться на одном листе.*

РЕФЕРАТ

**Выпускная квалификационная работа бакалавра:** 63 c., 16 рисун-ков, 6 таблиц, 10 источников, два приложения.

**Презентация:** 12 слайдов Microsoft PowerPoint.

НЕЙРОННЫЕ СЕТИ, БЕЗМАРКЕРНЫЙ ЗАХВАТ ДВИЖЕНИЯ, ОБРАБОТКА ИЗОБРАЖЕНИЙ, АНИМАЦИЯ, НЕЙРОСЕТЕВОЙ АНАЛИЗ ДАННЫХ

Работа посвящена исследованию погрешности численного моделирования процесса распространения электромагнитного импульса в однородном планарном волноводе.

Значительная часть работы уделена нахождению аналитического решения для двух способов описания входного импульса, а также теоретическому и экспериментальному исследованию оценки остатка ряда.

Часть работы связана с построением разностной схемы и исследованием погрешности численного решения с помощью вычислительного эксперимента.

В работе приведены результаты численного моделирования процесса распространения импульса в планарном волноводе.

Установлен факт существенного преимущества одного способа задания входного импульса над другим.

СОДЕРЖАНИЕ

[ВВЕДЕНИЕ 6](#_Toc164423847)

[1 Постановка задачи и цель работы 8](#_Toc164423848)

[2 Математическая основа 9](#_Toc164423849)

[3 Архитектура решения 10](#_Toc164423850)

[4 Методология эксперимента 11](#_Toc164423851)

[5 Результаты и сравнительный анализ с существующими моделями 12](#_Toc164423852)

[ЗАКЛЮЧЕНИЕ 13](#_Toc164423853)

[СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ 14](#_Toc164423854)

[ПРИЛОЖЕНИЕ А Моделирование распространения импульса в среде с переменным коэффициентом преломления 15](#_Toc164423855)

[ПРИЛОЖЕНИЕ Б Фрагменты кода программы 16](#_Toc164423856)

[Б.1 Функция построения аналитического решения для кусочно-гладкого импульса 16](#_Toc164423857)

ВВЕДЕНИЕ

Современные исследования в области компьютерного зрения и машинного обучения актуальны благодаря широкому спектру применения этих технологий в различных индустриях, включая автоматизированное видеонаблюдение, интерактивные системы, спортивный анализ и реабилитацию после травм. Особенно значимыми становятся методы захвата и анализа движений человека, которые позволяют улучшить интерфейсы человеко-машинного взаимодействия и повысить точность биомеханических исследований.

Для анализа движений человека на видео часто используются алгоритмы машинного обучения, которые могут автоматически распознавать и классифицировать различные типы движений из видеоданных. Основным преимуществом этих алгоритмов является способность обучаться на больших объемах данных и адаптироваться к новым, ранее неизвестным условиям, что делает их идеально подходящими для задач компьютерного зрения.

В данной работе рассматривается задача разработки и апробации программного средства, основанного на методах машинного обучения, для захвата и анализа движения человека в видеопотоке.

Данная работа содержит шесть разделов.

В первом разделе рассматривается постановка задачи, определяются начальные и краевые условия.

Во втором разделе осуществляется аналитическое решение поставленной задачи для волнового уравнения в виде бесконечного ряда Фурье при двух способах описания входного импульса.

В третьем разделе проводится исследование погрешности приближенно-аналитического решения, осуществляется оценка остатка ряда, а также экспериментальное исследование ее качества.

В четвертом разделе приведено численное решение поставленной задачи с помощью явной конечно-разностной схемы [3, 4].

В пятом разделе проводится сравнение полученных решений для двух способов задания входного импульса по критерию вычислительной трудоемкости.

И, наконец, в шестом разделе приведено описание программы, при помощи которой производится расчет решений и исследование их погрешности.

1. Постановка задачи и цель работы
   1. Описание задачи захвата движения в общем смысле

При анимации видеоигр, фильмов и создании контента для виртуальной реальности, технологии захвата движения играют ключевую роль, добавляя реализм и интерактивность. В спортивных науках, анализ видеозаписей тренировок и соревнований позволяет тренерам и спортсменам улучшать технику и стратегии. В медицине, особенно в реабилитации, захват движений используется для оценки и коррекции походки пациентов, что критически важно для успешного восстановления.

Одним из значительных преимуществ современных технологий захвата движения является их способность работать без маркеров, т.е. без необходимости носить специальные костюмы с датчиками. Это существенно упрощает процесс и делает технологию доступной для широкого круга пользователей, включая домашние условия. Также это позволяет анализировать естественные движения в реальной обстановке без искажения, вызванного дополнительным снаряжением.

Таким образом, способность точно анализировать человеческое поведение на видео открывает новые перспективы во множестве приложений, начиная от развлекательной индустрии до систем безопасности, где важно в реальном времени распознавать и реагировать на определенные действия и ситуации.

* 1. Задача захвата движения с точки зрения математики

Задача определения ключевых точек человеческого тела на видео может быть сформулирована как задача компьютерного зрения, в которой необходимо определить координаты предварительно определенных анатомических меток, таких как суставы, на изображении или в видеоряде. Математически это можно представить следующим образом.

Пусть обозначает кадр из видеопотока, а – набор ключевых точек, которые необходимо определить. Каждая точка описывается своими координатами на изображении . Задача алгоритма машинного обучения – максимизировать вероятность правильного определения этих координат, основываясь на обучающем наборе данных, содержащем аннотированные изображения.

Процесс обучения модели заключается в минимизации функции потерь, которая оценивает разницу между предсказанными алгоритмом координатами и истинными координатами точек на обучающих данных. Одним из популярных выборов для этой функции потерь является сумма квадратов разностей между предсказанными и истинными значениями координат:

*,*

где и – предсказанные координаты точек, а и – истинные координаты.

* 1. Цель работы

Цель данной работы — разработать, протестировать и исследовать применимость модели машинного обучения, способную с высокой точностью и в реальном времени определять ключевые точки человеческого тела на видео. Основные задачи включают улучшение точности захвата движения при различных условиях освещения, оптимизацию алгоритмов для работы в реальном времени на стандартном оборудовании и минимизацию нужды в ручной корректировке и аннотировании больших объемов данных.

Конечная цель исследования — не только создать эффективную техническую систему, но и продемонстрировать, как такие технологии могут быть интегрированы в реальные прикладные области, предоставляя значимую пользу в медицинских, спортивных и развлекательных приложениях.

1. Математическая основа модели
   1. Общий подход и архитектура модели
      1. Backbone сети

Backbone сеть в архитектуре Keypoint R-CNN является фундаментальным компонентом, который отвечает за первичное извлечение признаков из входных изображений. В контексте глубокого обучения, backbone — это обычно предварительно обученная сверточная нейронная сеть (CNN), которая преобразует исходные изображения в сложный набор признаков или карт признаков. Эти карты признаков служат основой для всех последующих этапов анализа и обработки, таких как детекция объектов, классификация и определение ключевых точек.

Распространенные архитектуры, используемые в качестве backbone в задачах компьютерного зрения, включают [Res2Net: A New Multi-scale Backbone Architecture]:

* ResNet (Residual Networks): Эта сеть использует так называемые "остаточные блоки", которые помогают обучать очень глубокие нейронные сети. Основное математическое выражение для слоя в ResNet выглядит следующим образом: , где – входной вектор признаков на уровне , – остаточная функция, а – веса, которые необходимо обучить.
* VGG (Visual Geometry Group): Простая и мощная архитектура, основанная на повторении блоков, состоящих из сверточных слоев с маленьким размером фильтра (3x3), за которыми следуют слои пулинга. В VGG все слои используют одинаковый шаг и дополнение нулями, что позволяет сохранять пространственные размеры через слои.

Feature Pyramid Network (FPN) является дополнением к стандартному backbone и предназначена для улучшения детекции объектов различного размера. FPN создает иерархию признаков на разных уровнях разрешения, что позволяет модели лучше адаптироваться к объектам разных масштабов. В FPN каждый уровень пирамиды создается путем слияния информации из двух источников: нижнего уровня с высоким разрешением и верхнего уровня с более глубокой семантической информацией. Это достигается с помощью операций свертки и апсемплинга [Effective Fusion Factor in FPN for Tiny Object Detection]:

,

где – признаки на уровне пирамиды, – выходные данные сверточного слоя на уровне , и – операция апсемплинга [Guided Upsampling Network for Real-Time Semantic Segmentation].

* + 1. RoI Pooling

RoI (Region of Interest) Pooling — это техника, используемая в сверточных нейронных сетях для извлечения фиксированного размера признаков из произвольно размерных регионов. Эта операция критически важна в задачах, где необходимо обработать локальные области изображения, такие как детекция объектов и сегментация. RoI Pooling был популяризирован архитектурами, такими как Fast R-CNN, для улучшения производительности моделей в этих задачах.

Рассмотрим входную карту признаков размером , где и обозначают высоту, ширину и глубину карты признаков соответственно. Пусть дано регионов интереса, каждый из которых задан четырьмя координатами: , где и обозначают верхний левый и нижний правый углы региона на карте признаков.

Для каждого региона RoI Pooling операция выполняется следующим образом [ROI Pooled Correlation Filters for Visual Tracking]:

1. Преобразование размеров: Регион преобразуется к фиксированному размеру , который задается заранее. Это делается для стандартизации выходных данных, позволяя использовать их на следующих слоях нейронной сети.
2. Разделение на ячейки: Преобразованный регион делится на ячеек равного размера.
3. Пулинг: В каждой ячейке выполняется операция пулинга для получения одного значения признака. Пулинг [A Theoretical Analysis of Feature Pooling in Visual Recognition] определяется, как:

,

где – это значение признака в ячейке , а – множество пикселей в ячейке входной карты признаков.

* + 1. Классификационные и регрессионные слои, блок ключевых точек

После извлечения признаков с помощью RoI Pooling, следующий шаг заключается в использовании классификационных и регрессионных слоев для определения класса объектов и точного позиционирования их ограничивающих рамок. Классификационный слой применяет функцию softmax для оценки вероятности принадлежности каждого предложенного региона к одному из возможных классов объектов:

,

где – вероятность класса для региона , – сырые оценки классификации для класса , а – общее количество классов.

Регрессионный слой затем адаптирует ограничивающие рамки, предложенные RPN, для того чтобы они как можно точнее обрамляли детектируемый объект. Это достигается путем корректировки четырех параметров рамки: центра , ширины , ширины и высоты . Преобразование координат рамки выражается через:

,

где - предсказания регрессии относительно исходной рамки [A High-Speed and Low-Complexity Architecture for Softmax Function in Deep Learning].

После классификации объектов и уточнения их рамок, выполняется задача определения ключевых точек. Каждая детектированная рамка объекта снабжается набором признаков, из которых затем предсказываются ключевые точки. Это достигается с помощью регрессии для каждой точки внутри рамки:

,

где – координаты предсказанной ключевой точки, – исходные координаты точки в рамке, а – смещения, предсказанные сетью [Local keypoint-based Faster R-CNN].

1. Архитектура решения
2. Методология эксперимента
3. Результаты и сравнительный анализ с существующими моделями

ЗАКЛЮЧЕНИЕ

В ходе работы было проведено исследование погрешности моделирования распространения электромагнитного импульса в планарном волноводе.

С помощью метода разделения переменных найдено решение краевой задачи в виде бесконечных рядов Фурье для двух способов описания входного импульса.

Для контроля погрешности усечения ряда была получена оценка его остатка. Было установлено, что получаемое количество суммируемых элементов при заданной точности является существенно избыточным. Поэтому была предложена методика экспериментального определения количества суммируемых членов ряда, которая позволяет уменьшить это количество для кусочно-гладкого импульса минимум в пять раз (а при отсечении фронтального участка в 11000 раз), для гладкого – в 350−2400 раз, и при этом гарантировать обеспечение требуемой точности.

Также было получено численное решение поставленной задачи с помощью метода конечных разностей. Проведено исследование погрешности разностного решения, в результате которого было замечено, что при использовании способа описания входного импульса кусочно-гладкой функцией скорость сходимости разностной схемы гораздо меньше, чем при использовании способа описания импульса гладкой функцией.

Разработана программа для численного моделирования процесса распространения электромагнитного импульса на основе полученных решений в однородной среде и при размещении в волноводе оптического клина.

Проведено сравнение полученных решений для двух способов задания входного импульса по критерию вычислительной трудоемкости. Установлено, что описание входного импульса гладкой функцией дает экономию по времени построения решения в 3060 раз при использовании аналитического метода решения и в 190 раз при использовании разностного метода.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Козлова, Е.С. Моделирование распространения короткого двумерного импульса света / Е.С. Козлова, В.В. Котляр // Компьютерная оптика. − 2012. − Т. 36, № 2. − C. 158-164.
2. Козлова, Е.С. Моделирование предвестников Зоммерфельда и Бриллюэна в среде с частотной дисперсией на основе разностного решения волнового уравнения / Е.С. Козлова, В.В. Котляр // Компьютерная оптика. − 2013. − Т. 37, № 2. − C. 146-154.
3. Самарский, А.А. Теория разностных схем : учеб. пособие / А.А. Самарский. − Москва : Наука, 1977. − 656 с.
4. Самарский, А.А. Введение в численные методы : учеб. пособие для вузов / А.А. Самарский. − Москва : Наука, 1987. − 288 с.
5. Евтихов, М.Г. Соотношения Снеллиуса и Френеля для электромагнитных волн с постоянной линейной поляризацией // Радиотехника и электроника. − 2010. − Т. 55, № 8. − C. 915-922.
6. Тихонов, А.Н. Уравнения математической физики : учеб. пособие / А.Н. Тихонов, А.А. Самарский. − Москва : Наука, 1972. − 736 с.
7. Зайцев, В.Ф. Метод разделения переменных в математической физике : учеб. пособие / В.Ф. Зайцев, А.Д. Полянин. − Санкт-Петербург : Книжный Дом, 2009. − 92 с.
8. Владимиров, В.С. Уравнения математической физики : учеб. для вузов / В.С. Владимиров, В.В. Жаринов. − Москва : Физматлит, 2004. − 400 с.
9. Волков, И.К. Случайные процессы : учеб. для вузов / И.К. Волков, С.М. Зуев, Г.М. Цветкова. − Москва : Изд-во МГТУ им. Баумана, 1999. − 448 с.
10. Фихтенгольц, Г.М. Курс дифференциального и интегрального исчисления. В 3 т. Т. 3 / Г.М. Фихтенгольц. – Москва : Физматлит, 2001. – 662 с.

ПРИЛОЖЕНИЕ А  
Моделирование распространения импульса в среде   
с переменным коэффициентом преломления

Предположим, что на пути распространения импульса размещен оптический клин из материала, имеющего более высокий коэффициент преломления. Схематическое изображение такого волновода показано на рисунке А.1.

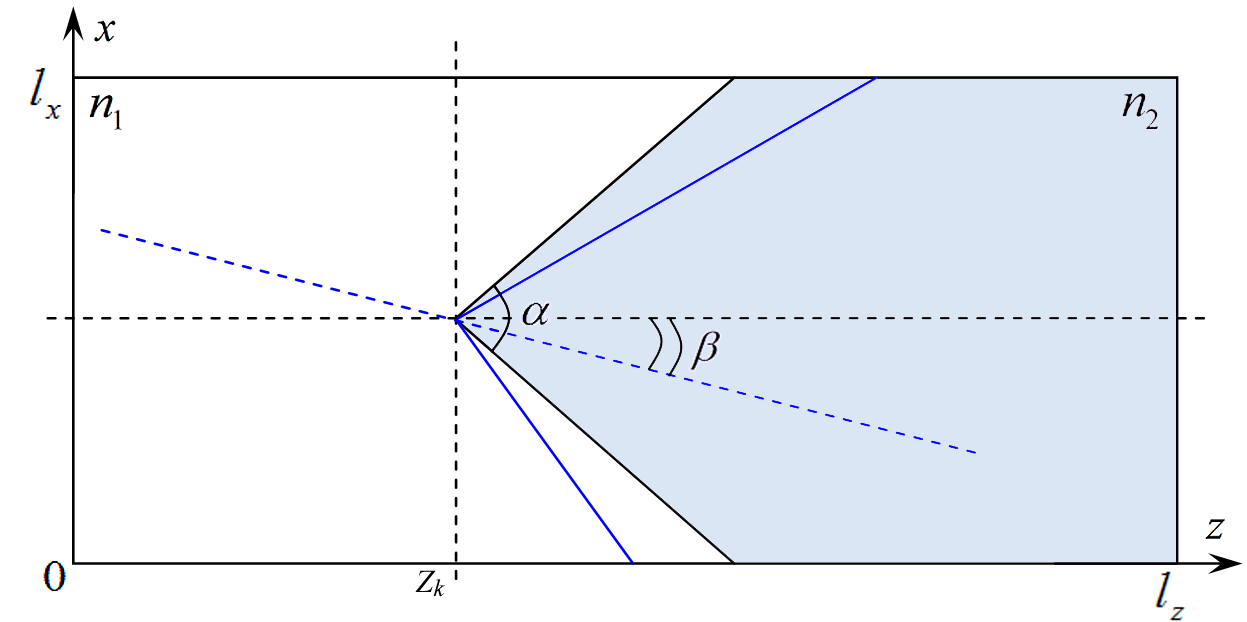


Рисунок А.1 – Изображение волновода с клином

Разработанная и оттестированная программа была применена для расчета процесса прохождения импульса через такой клин.

Проведем сравнение разностных решений для каждого случая. Зафиксируем:  Задавая уровень погрешности , будем изменять сетку таким образом, чтобы схема была устойчивой и фактическая погрешность практически совпадала с . Результаты приведены в таблице А.1.

Таблица А.1 – Сравнение разностных решений

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | 0,02 | 0,01 | 0,004 |
| Количество узлов сетки | Кусочно-гладкий импульс | 4090601 | 64682001 | 1029287201 |
| Гладкий импульс | 17576 | 262701 | 4090601 |
| Разница (во сколько раз) | 232,7 | 246,2 | 251,6 |
| Время выполнения, мс | Кусочно-гладкий импульс | 761 | 12107 | 135537 |
| Гладкий импульс | 5 | 50 | 766 |
| Разница (во сколько раз) | 152,2 | 242,1 | 176,9 |

ПРИЛОЖЕНИЕ Б  
Фрагменты кода программы

Б.1 Функция построения аналитического решения для кусочно-гладкого импульса

public double[] getAnalyticalResult1()

{

double[] res = new double[G + 1];

double sum = 0;

double z = 0;

double k = 0;

double wm = 0;

double sinwmt = 0;

double coswmt = 0;

double w = (2 \* Math.PI \* c) / lambda;

double w2 = w \* w;

double[] wm2 = new double[count + 1];

double x = lx / 2;

double a2 = (eps \* w2) / (c \* c) - (Math.PI \* Math.PI) / (lx \* lx);

double b2 = (c \* c \* a2) / eps;

double sinwt = Math.Sin((2 \* Math.PI \* tmkm \* Math.Sqrt(eps)) / lambda);

for (int m = 0; m <= count; m++)

{

wm2[m] = ((c \* c \* Math.PI \* Math.PI) / eps) \* ((1 / (lx \* lx)) + (((1 + 2 \* m) \*

(1 + 2 \* m)) / (4 \* lz \* lz)));

}

if (tmkm <= tmkm1)

{

for (int g = 0; g <= G; g++)

{

z = hg \* g;

sum = 0;

k = 0;

sinwmt = 0;

wm = 0;

for (int m = 0; m <= count; m++)

{

wm = Math.Sqrt(wm2[m]);

sinwmt = Math.Sin(Math.Sqrt(wm2[m]) \* tmkm \* Math.Sqrt(eps) / c);

k = (b2 / (wm2[m] - w2)) \* (wm \* sinwt - w \* sinwmt) - w \* sinwmt;

sum += Math.Sin(Math.PI \* x / lx) \* Math.Sin(Math.PI \* (1 + 2 \* m) \* z / (2

\* lz)) \* k\*(1-Math.Cos(Math.PI\*(1 + 2 \*m)/2))\*(4/(wm\*Math.PI\*(1 + 2 \* m)));

}

res[g] = sum + sinwt \* Math.Sin(Math.PI \* x / lx);

}

}

else

{

for (int g = 0; g <= G; g++)

{

z = hg \* g;

sum = 0;

k = 0;

sinwmt = 0;

coswmt = 0;