ФИЗИЧЕСКИЕ ОСНОВЫ ПРИБОРОСТРОЕНИЯ

ISSN: 2225-4293
Online ISSN: 2712-7885

2020. Т.9. №2 (36)

ФИЗИЧЕСКИЕ ОСНОВЫ ПРИБОРОСТРОЕНИЯ. 2020. Т.9. №2 (36)

Научно-технологический центр уникального
приборостроения Российской академии наук

Учредители: Научно-технологический центр уникального приборостроения Российской академии наук, Кравченко Виктор Филиппович, Пустовойт Владислав Иванович

Издатель: Федеральное государственное бюджетное учреждение науки Научно-технологический центр уникального приборостроения Российской академии наук

Журнал зарегистрирован 15 февраля 2000 г. Министерством Российской Федерации по делам печати, телерадиовещания и средств массовой информации

Свидетельство о регистрации ПИ № 77-1685

Физические Основы Приборостроения

2020. Т.9. №2 (36)

ISSN 2225-4293

doi: 10.25210/jfop-2002

Online ISSN 2712-7885

Журнал «Физические основы приборостроения» включен:

- В **Российский индекс научного цитирования (РИНЦ)**.
- В список журналов базы данных **Russian Science Citation Index (RSCI)** на платформе **Web of Science**.

• В **Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук.**

РЕДКОЛЛЕГИЯ

Пустовойт В.И., гл. редактор, академик РАН, д.ф.-м.н., проф.
Кравченко В.Ф., зам. гл. редактора, д.ф.-м.н., проф.
Боголюбов А.Н., д.ф.-м.н., проф.
Боритко С.В., д.ф.-м.н., проф.
Быков А.А., д.ф.-м.н.
Булатов М.Ф., д.ф.-м.н., проф.
Виноградов Е.А., член-корр. РАН, д.ф.-м.н., проф.
Гуляев Ю.В., академик РАН, д.ф.-м.н., проф.
Дианов Е.М., академик РАН, д.ф.-м.н., проф.
Дмитриев А.С., д.ф.-м.н., проф.
Крохин О.Н., академик РАН, д.ф.-м.н., проф.
Крюковский А.С., д.ф.-м.н., проф.
Кутуза Б.Г., д.ф.-м.н., проф.
Лукин Д.С., д.ф.-м.н., проф.
Мазур М.М., д.т.н.
Морозов А.Н., член-корр. РАН, д.ф.-м.н., проф.
Назаров Л.Е., д.ф.-м.н.
Панич А.Е., д.т.н., проф.
Панич А.А., д.т.н., проф.
Садовничий В.А., академик РАН, д.ф.-м.н., проф.
Самохин А.Б., д.ф.-м.н., проф.
Сидняев Н.И., д.т.н., проф.
Синявский Г.П., д.ф.-м.н., проф.
Татаренко Н.И., д.ф.-м.н.
Федоров И.Б., академик РАН, д.т.н., проф.

ЗАРУБЕЖНЫЕ ЧЛЕНЫ РЕДКОЛЛЕГИИ

Буц В.А., д.ф.-м.н., проф. (Украина)
Волосюк В.К., д.т.н., проф. (Украина)
Загородний А.Г., академик НАН Украины, иностранный член РАН, д.ф.-м.н., профессор (Украина)
Кобаяши К., проф. (Япония)
Кураев А.А., д.ф.-м.н., проф. (Беларусь)
Луценко В.И., д.ф.-м.н., (Украина)
Мартirosян Р.М., академик НАН Республики Армения, иностранный член РАН, д.ф.-м.н., проф. (Республика Армения)
Мележик П.Н., академик НАН Украины, д.ф.-м.н., проф. (Украина)
Миттра Р., проф. (США)
Павликов В.В., д.т.н., (Украина)
Перес-Меана Э., проф. (Мексика)
Пономарев В.И., д.т.н., проф. (Мексика)
Ра Д., проф. (Корея)
Хашимото М., проф. (Япония)
Яковенко В.М., академик НАН Украины, д.ф.-м.н., проф. (Украина)

EDITORIAL BOARD

Pustovoit, V.I., Editor-in-Chief, Academician of RAS, Dr. Sci. (Phys.- Math.), Prof.
Kravchenko, V.F., Deputy Editor-in-Chief, Dr. Sci. (Phys.- Math.), Prof.
Bogolyubov, A.N., Dr. Sci. (Phys.- Math.), Prof.
Boritko, S.V., Dr. Sci. (Phys.- Math.), Prof.
Bykov, A.A., Dr. Sci. (Phys.- Math.)
Bulatov, M.F., Dr. Sci. (Phys.- Math.), Prof.
Vinogradov, E.A., Corresponding Member of RAS, Dr. Sci. (Phys.- Math.), Prof.
Gulyaev, Yu.V., Academician of RAS, Dr. Sci. (Phys.- Math.), Prof.
Dianov, E.M., Academician of RAS, Dr. Sci. (Phys.- Math.), Prof.
Dmitriev, A.S., Dr. Sci. (Phys.- Math.), Prof.
Krohin, O.N., Academician of RAS, Dr. Sci. (Phys.- Math.), Prof.
Krukovsky, A.S., Dr. Sci. (Phys.- Math.), Prof.
Kutuza, B.U., Dr. Sci. (Phys.- Math.), Prof.
Lukin, D.S., Dr. Sci. (Phys.- Math.), Prof.
Mazur, M. M., Dr. Sci. (Techn.)
Morozov, A. N., Corresponding member of RAS, Dr. Sci. (Phys.- Math.), Prof.
Nazarov, L.E., Dr. Sci. (Phys.- Math.)
Panich, A.E., Dr. Sci. (Techn.), Prof.
Panich, A.A., Dr. Sci. (Techn.), Prof.
Sadovnichiy, V.A., Academician of RAS, Dr. Sci. (Phys.- Math.), Prof.
Samokhin, A.B., Dr. Sci. (Phys.- Math.), Prof.
Sidnyayev, N.I., Dr. Sci. (Techn.), Prof.
Synyavskiy, G.P., Dr. Sci. (Phys.- Math.), Prof.
Tatarenko, N.I., Dr. Sci. (Phys.- Math.)
Fedorov, I.B., Academician of RAS, Dr. Sci. (Techn.), Prof.

FOREIGN EDITORIAL BOARD MEMBERS

Buts, V.A., Dr. Sci. (Phys.- Math.), Prof. (Ukraine)
Volosyuk, V. K., Dr. Sci. (Techn.), Prof. (Ukraine)
Zagorodny, A.G., Academician NAS of Ukraine, Foreign Member of the RAS, Dr. Sci. (Phys.- Math.), Prof. (Ukraine)
Kobayashi, K., Prof. (Japan)
Kuraev, A.A., Dr. Sci. (Phys.- Math.), Prof. (Belarus)
Lutsenko, V.I., Dr. Sci. (Phys.- Math.) (Ukraine)
Martirosyan, R.M., Academician of NAS of Republic of Armenia, Foreign Member of the RAS, Dr. Sci. (Phys.- Math.), Prof. (Republic of Armenia)
Melezhik, P.N., Academician of NAS of Ukraine, Dr. Sci. (Phys.- Math.), Prof. (Ukraine)
Mittra, R., Prof. (USA)
Pavlikov, V.V., Dr. Sci. (Techn.) (Ukraine)
Perez-Meana, H., Prof. (Mexico)
Ponomaryov, V.I., Dr. Sci. (Techn.), Prof. (Mexico)
Ra, J-W., Prof. (Korea)
Hashimoto, M., Prof. (Japan)
Yakovenko, V.M., Academician NAS of Ukraine, Dr. Sci. (Phys.- Math.), Prof. (Ukraine)

СОДЕРЖАНИЕ

CONTENTS

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ФИЗИЧЕСКИХ ПРОЦЕССОВ

MATHEMATICAL MODELING OF PHYSICAL PROCESSES

Белов А.А., Боголюбов А.Н.,
Домбровская Ж.О., Жбанников С.О.
Сверхбыстрый метод расчёта
одномерных задач фотоники

2

Belov, A.A., Bogolyubov, A.N.,
Dombrovskaya, Zh.O., and Zhbannikov, S.O.
Superfast Method for One-Dimensional
Photonic Problems Computation

Назаров Л.Е.

Помехоустойчивость оптимального посимвольного приема
сигналов в недвоичных полях Галуа

10

Nazarov, L.E.

The Noise-Immunity of Optimal Symbol-by-Symbol Decoding
Algorithm in Non-Binary Galua Fields

Ерофеенко В.Т., Урбанович А.И.

Начально-краевая задача экранирования импульсных
электромагнитных полей экранами из пермаллоя для
численного моделирования

16

Erofeenko, V.T., Urbanovich, A.I.

Initial Boundary-Value Problem of Screening Pulsed
Electromagnetic Fields by Permalloy Shields
for Numericall Modeling

Егоров Д.П., Кравченко О.В.,
Митрофанова А.Ю., Чуриков Д.В.

Генератор тестовых изображений для детекторов углов

28

Egorov, D. P., Kravchenko, O. V.,

Mitrofanova, A. Yu., and Churikov, D. V.
Image Generator for Testing Corner Detectors

НОВЫЕ МАТЕМАТИЧЕСКИЕ И ФИЗИЧЕСКИЕ МЕТОДЫ

NEW MATHEMATICAL AND PHYSICAL METHODS

Болотова Е.Е., Бутенко Ю.И., Сидняев Н.И.
Язык логики предикатов в системах обработки
информации в базах знаний

37

Bolotova, E. E., Butenko, Yu. I., and Sidnyaev, N. I.
Predicate Logic in Information Processing Systems
in Knowledge Bases

Коновалов Я.Ю.

Новое семейство бесконечно
гладких функций с компактным носителем

48

Konovalov, Ya.Yu.

New Family of Infinitely
Smooth Compactly Supported Functions

ОПТИЧЕСКИЕ МАТЕРИАЛЫ И ТЕХНОЛОГИИ

OPTICAL MATERIALS AND TECHNOLOGIES

Петров Н.И., Соколов Ю.М., Хромов М.Н.
3Д дисплеи без очков на основе проекторов
и мобильных телефонов

60

Petrov, N.I., Sokolov, Y.M., and Khromov, M.N.
Glassless 3D Imaging Systems Based
on Projectors and Mobile Phones

МЕТОДЫ РАДИОЛОКАЦИОННЫХ И РАДИОМЕТРИЧЕСКИХ ИЗМЕРЕНИЙ

METHODS OF RADAR AND RADIOMETRIC MEASUREMENTS

Лаговский Б.А., Самохин А.Б.

Особенности решений задач радиолокации
при использовании сверхширокополосных сигналов

74

Lagovsky, B.A., Samokhin, A.B.

Features of Solving Radar Problems
Using Ultra-Wide-Band Signals

ЭЛЕКТРОДИНАМИКА ИСКУССТВЕННЫХ СРЕД И СТРУКТУР

ELECTRODYNAMICS OF ARTIFICIAL MATERIALS AND STRUCTURES

Буц В. А., Тарасов Д. В.
Об условиях взаимодействия волн

82

Buts, V. A., Tarasov, D. V.

About Conditions of Wave Interaction

Рубрики журнала

91

Journal Headings

Правила оформления рукописей

92

Rules for the Preparation of Manuscripts

doi: 10.25210/jfop-2002

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ФИЗИЧЕСКИХ ПРОЦЕССОВ

УДК 519.634

СВЕРХБЫСТРЫЙ МЕТОД РАСЧЁТА ОДНОМЕРНЫХ ЗАДАЧ ФОТОНИКИ

© Авторы, 2020

doi: 10.25210/jfor-2002-002009

Белов А. А. — к.ф.м.н., сотрудник кафедры математики физического факультета Московского государственного университета имени М.В. Ломоносова. E-mail: aa.belov@physics.msu.ru

Боголюбов А. Н. — д.ф.м.н., профессор кафедры математики физического факультета Московского государственного университета имени М.В. Ломоносова. E-mail: bogan7@yandex.ru

Домбровская Ж. О. — к.ф.м.н., сотрудник кафедры математики физического факультета Московского государственного университета имени М.В. Ломоносова. E-mail: dombrovskaya@physics.msu.ru

Жбанников С. О. — студент кафедры математики физического факультета Московского государственного университета имени М.В. Ломоносова. E-mail: zhbannikovstas@yandex.ru

Аннотация

Предложен новый экономичный численный метод для расчета распространения электромагнитных волн через одномерный фотонный кристалл. Закон частотной дисперсии материалов структуры может быть произвольным. Приведены примеры тестовых и прикладных расчетов.

Ключевые слова: матричный метод, преобразование Фурье, фотонный кристалл

Abstract

New efficient numerical method for calculating electromagnetic wave propagation through one-dimensional photonic crystal is proposed. Material dispersion can be arbitrarily specified. Examples of test and applied computations are given.

Keywords: matrix method, Fourier transform, photonic crystal

Введение

В настоящее время большой теоретический и практический интерес вызывает применение метаматериалов — искусственно структурированных композитных материалов, сильно взаимодействующих с электромагнитными волнами [1–3]. В частности, большой интерес представляет исследования электродинамических систем на основе фотонных кристаллов. Интенсивно изучаются свойства фотонно-кристаллических структур, разрабатываются методы их исследования, создаются и исследуются различные устройства на таких структурах, реализуются на практике теоретически предсказанные эффекты в фотонных кристаллах. Интерес к таким задачам обусловлен, прежде всего, тем, что фотонные кристаллы позволяют строить приборы и устройства, обладающие уникальными свойствами.

Фотонный кристалл — это материал, структура которого характеризуется периодическим изменением коэффициента преломления. Рассматриваются одномерные, двумерные и трехмерные фотонные кристаллы. Как известно, кристаллы всех типов могут эффективно рассеивать некоторые виды излучения при условии, что параметры решетки кристалла имеют тот же порядок, что и длина волны излучения. При прохождении электромагнитной волны через одномерный фотонный кристалл возникают многократные переотражения от границ слоев. Интерференция переотраженных волн приводит к появлению «фотонных запрещенных зон» (photonic band gap, PBG) в спектрах отражения и прохождения. Если запрещенные зоны для всех направлений в кристалле перекрываются, запрещенная зона является полной, что означает, что в определенном спектральном диапазоне свет любой поляризации не может распространяться в кристалле ни в каком направлении в одном, двух или трех измерениях.

Такие свойства делают фотонно-кристаллические структуры весьма интересными своими приложениями, например, следующие:

- волноводы на основе фотонных кристаллов могут быть очень компактны и обладать малыми потерями;
- фотонные сверхпроводники способны работать с большими частотами и отлично совмещаются с фотонными изоляторами и полупроводниками;
- благодаря упорядоченному характеру явления захвата в фотонном кристалле, на основе этих сред возможно построение оптических запоминающих и логических устройств;
- при наклонном падении волны на фотонный кристалл возможно формирование внутри него связанного состояния, распространяющегося вдоль его поверхности; такие связанные состояния широко используются в интегральной фотонике в качестве носителя сигнала.

При исследовании свойств фотонно-кристаллических структур и приборов и устройств на их основе весьма широко применяются методы математического моделирования [3]. В частности, используются методы конечных разностей и конечных элементов, а также матричные методы. Эти методы хорошо зарекомендовали себя при решении стационарных задач. Однако большинство задач, связанных с математическим моделированием систем и устройств на фотонных кристаллах, являются нестационарными. Простейшим примером такой задачи является задача о падении короткого электромагнитного импульса на поверхность фотонного кристалла.

Наиболее широко применяемым методом для решения таких задач является метод FDTD — метод конечных разностей во временной области, в основе которого лежит предложенная в 1966 году К. Йе оригинальная конечно-разностная схема для решения уравнений Максвелла [4]. Хотя в настоящее время алгоритм Йе широко используется в электродинамике [5], его применение при расчете одномерных фотонных кристаллов связано со значительными сложностями [6].

В данной работе на основе стационарных матричных методов [7] и преобразования Фурье построен нестационарный матричный метод для расчета распространения излучения через одномерный фотонный кристалл.

Нестационарный матричный метод

Основная идея [8] предлагаемого нестационарного метода заключается в следующем. Поскольку рассматривается линейная задача расчета одномерного фотонного кристалла, то для неё справедлив принцип суперпозиции. Выполним Фурье-преобразование падающего излучения с помощью квадратурной формулы трапеций. Представим падающий импульс в виде суперпозиции плоских монохроматических волн, применяя численное преобразование Фурье и вычисляя интегралы с помощью формулы трапеции. Для каждой из плоских монохроматических волн решим стационарную задачу о падении такой волны на фотонный кристалл, применяя стандартный стационарный матричный метод, использующий метод матриц Беремана [7], найдем амплитудные коэффициенты отражения и прохождения. Затем, выполняя численно обратное преобразование Фурье, также применяя формулу трапеций, найдем временные развертки прошедшего и отраженного полей

$$E(z, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} E(\omega) e^{\frac{i\omega}{c}z - i\omega t} d\omega \cong \frac{1}{\sqrt{2\pi}} \sum_{n=1}^{N+1} g_n E_0(\omega_n) e^{\frac{i\omega_n}{c}z - i\omega_n t} h_{\omega_n},$$

где

$$g_n = \frac{1}{2}, \quad n = 1, N+1; \quad g_n = 1, \quad n \neq 1, N+1.$$

Предложенный метод имеет ряд преимуществ по сравнению с известными методами решения нестационарных задач.

Во-первых, он применим не только для нормального, но и для наклонного падения волны на поверхность фотонного кристалла (поскольку такая возможность есть в известных стационарных матричных методах). Для методов типа FDTD задача о наклонном падении на поверхность даже одномерного фотонного кристалла является двумерной, то есть имеет достаточно высокую вычислительную трудоемкость.

Во-вторых, по построению, метод Фурье-преобразования позволяет учитывать произвольный закон дисперсии материальных параметров (в том числе, заданный таблично).

В-третьих, если временная развертка падающей волны является локализованной, то квадратурные формулы, реализующие преобразование Фурье, сходятся не степенным, а экспоненциальным образом [10]. Поэтому предложенный метод является сверхбыстрым.

Для контроля фактической точности квадратур использовался метод сгущающихся сеток и апостериорные оценки, предложенные в [11].

Пусть имеется световой импульс с известным спектром $E_\lambda(\lambda)$. С помощью замены $\lambda = \frac{2\pi c}{\omega}$ перейдем к частотному представлению $E_\omega(\omega)$.

Зададим равномерную сетку по частоте $\Omega_N = \{\omega_n; n = 1, \dots, N\}$ и сопоставим ей соответствующую неравномерную сетку по длине волны $\Lambda_N = \{\lambda_n; n = 1, \dots, N\}$. Введем сеточную функцию $E_n = E_\omega(\omega_n) = E_\lambda(\lambda_n)$.

Для каждого узла решим стационарную задачу поиска коэффициентов отражения слоистой структуры. После чего вычислим спектр отраженного импульса на изначальной частотной сетке

$$E_n^r = r_n E_n. \quad (1)$$

Чтобы найти пространственное распределение падающего и отражённого импульсов воспользуемся обратным фурье-преобразованием

$$E(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} E(w) e^{-iwt} dw \approx \frac{1}{\sqrt{2\pi}} \int_a^b E(w) e^{-iwt} dw, \quad (2)$$

где $E(\tau)$ — нормированный на единицу начальный импульс на временной сетке; $|\bar{E}(\tau)|$ — максимальная амплитуда начального импульса

$$E^r(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} E^r(w) e^{-iwt} dw \approx \frac{1}{\sqrt{2\pi}} \int_a^b E^r(w) e^{-iwt} dw. \quad (3)$$

Второй знак равенства в обоих случаях обеспечивается выбором достаточно большого по сравнению с шириной спектра импульса частотного интервала, так как это позволяет считать спектр импульса локализованным на интервале $[a, b]$.

Ограничение отрезка интегрирования позволяет использовать равномерные сетки. Известно [9,10], что, если (а) нечетные производные подынтегральной функции достаточно малы на границах интервала и (б) сетки равномерные, то формула трапеций имеет сверхстепенную скорость сходимости. Приведем доказательство этого факта.

Обозначим подынтегральную функцию как $U(w) = E(w) e^{-i\omega t}$. Зададим равномерную временную сетку $T_N = \{\tau_n; n = 1, \dots, N\}$ с N узлами и разложим точное решение в произвольной точке временной сетки $E(\tau_n)$ по формуле Маклорена

$$E(\tau_n) = d\omega \left(\frac{U(\omega_0)}{2} + U(\omega_1) + \dots + U(\omega_{N-1}) + \frac{U(\omega_N)}{2} \right) + \sum_{p=1} (-1)^p a_p d\omega^{2p} (U^{(2p-1)}(\omega_N) - U^{(2p-1)}(\omega_0)), \quad (3)$$

где

$$a_p \sim M_{2p-1} = \max |U^{(2p-1)}(\omega)|.$$

Учтём, что начальный и граничный импульсы ограничены в спектре, соответственно при достаточно большом частотном интервале расчёта получаем

$$E^{(2p-1)}(a) \approx E^{(2p-1)}(b) \approx 0 \Rightarrow U^{(2p-1)}(\omega_N) - U^{(2p-1)}(\omega_0) \approx e^{-i\omega t} \cdot 0 = 0.$$

Следовательно,

$$E(\tau_n) = d\omega \left(\frac{U(\omega_0) + U(\omega_N)}{2} \right) + \sum_{k=1}^{N-1} U_n(\omega_k) + o(\Delta\omega^p) = E\tau_n + o(\Delta\omega^p), \forall_p. \quad (4)$$

Здесь $E\tau_n$ — численное решение интеграла Фурье в узле временной сетки τ_n .

Соответственно имеем сверхстепенную сходимость к точному решению $E\tau_n \rightarrow E(\tau_n)$ в каждой точке временной сетки T_N , при $N \rightarrow \infty$.

Рис. 1. Зависимость показателя преломления материала от длины волны для SiO_2 (а) и Ta_2O_5 (б).

Верификация предложенного алгоритма

Предлагаемый в работе алгоритм был реализован в виде рабочей программы. Правильность ее работы проверялась двумя способами.

Для проверки блока решения стационарной задачи проводились расчеты спектра отражения от фотонного кристалла. Пусть плоская монохроматическая волна наклонно падает из воздуха на структуру, включающую 7 бислоев SiO_2 и Ta_2O_5 толщиной 160 и 112 нм соответственно [13]. Поверх последнего слоя Ta_2O_5 располагается покрытие SiO_2 толщиной 260 нм. Угол наклона равен 45 градусов. Зависимость показателей преломления SiO_2 и Ta_2O_5 от частоты приведены на рис. 1.

Полученный спектр сравнивался со спектром, вычисленным с помощью бикомпактной разностной схемы [8]. Результаты обоих расчетов приведены на рис. 2. В спектре хорошо видна запрещенная зона, соответствующая практически полному отражению от фотонного кристалла.

Рис. 2. Результаты расчёта спектр отражения от $\text{SiO}_2/\text{Ta}_2\text{O}_5$ образца матричным методом (зелёная линия) и бикомпактной разностной схемой (чёрный пунктир).

Видно, что с графической точностью кривые совпадают. Различие между ними сводится к сдвигу по горизонтальной оси. Поэтому для определения количественного расхождения будем сравнивать положения минимумов. Такая процедура аналогична метрике Хаусдорфа, предложенной в [12] для сравнения решений жестких задач, имеющих участки чрезвычайно быстрого изменения. Расхождение кривых на рис. 2 в указанном смысле согласуется с апостериорной оценкой точности разностного расчета. Это подтверждает правильность реализации блока решения стационарной задачи.

Для верификации метода решения нестационарной задачи проводились расчеты тестовой задачи с известным точным решением. Рассмотрим наклонное падение заданного импульса на границу раздела однородных недиспергирующих диэлектриков. Точное решение этой задачи имеет вид

$$E_{\text{tot}} = E_0 \left(\frac{z}{c} - t \right) + \frac{1 - \sqrt{\varepsilon}}{1 + \sqrt{\varepsilon}} E_0 \left(-\frac{z}{c} - t + \frac{2b}{c} \right), \quad (6)$$

где b — расстояние от источника до структуры, $0 \leq z \leq b$.

Погрешность расчета определялась непосредственно как разность численного и точного решений. Полученные погрешности Δ в зависимости от числа узлов сеток по времени и по частоте приведены на рис. 3. Масштаб графика полулогарифмический. Видно, что скорость сходимости намного превышает степенную. Однако надежно установить ее закон на данном примере не удастся, поскольку погрешности быстро достигают ошибок округления.

Рис. 3. Относительная погрешность решения нестационарной задачи в зависимости от числа шагов сетки.

Проведенные расчеты являются достаточно убедительной верификацией построенных методов и наглядно иллюстрируют их преимущества по сравнению с известными подходами.

Блоховские волны

Рассмотрим задачу о прохождении излучения через ФК в нестационарной постановке. Пусть из SiO_2 на ФК наклонно падает гауссов импульс с центральной длиной волны 800 нм и полушириной на половине высоты 30 нм. Спектр этого импульса показан на рис. 4. Угол падения по-прежнему равен 45 градусов.

Рис. 4. Спектр падающего импульса.

На границе раздела « SiO_2 -воздух» возникает полное внутреннее отражение. В фотонном кристалле формируется связанное состояние (блоховская поверхностная волна). Эта волна распространяется вдоль поверхности фотонного кристалла, теряя энергию за счет излучения. Поэтому ее время жизни оказывается конечным. Напомним, что временем жизни такого связанного состояния называется время, за которое интенсивность свечения уменьшается в e раз. Спектр отражения описанной структуры представлен на рис. 5. Блоховской волне соответствует узкий минимум в этом спектре.

Рис. 5. Зависимость коэффициента отражения ФК от длины волны, при падении света из SiO_2 .

Известно, что время жизни поверхностной волны связано с шириной минимума в спектре отражения

$$\tau = \frac{1}{\Delta\omega} = \frac{1}{2\pi c \left(\frac{1}{\lambda_2} - \frac{1}{\lambda_1} \right)} = \frac{\lambda_0^2}{2\pi c \Delta\lambda}, \quad (7)$$

где λ_1 , λ_2 , λ_0 — длины волн на полуширине провала и резонансная длина волны соответственно.

Для данного фотонного кристалла $\lambda_1 = 829.1$, $\lambda_2 = 829.7$, $\lambda_0 = 829.4$ нм. Тогда по формуле (7) оценим $\tau \approx 600$ фс. Теперь получим более точное значение времени жизни поверхностной волны. Для начала воспользуемся построенным в данной работе нестационарным методом, чтобы найти характеристики отражённого импульса. По формуле (1) получим спектр отражённого импульса и, используя преобразования (2), (4), получим его развёртку на временной сетке, представленную на рис. 6.

Для проверки сходимости нестационарного матричного метода и точности полученного результата воспользуемся апостериорной оценкой точности предложенной в [11]. Построим решения на серии сгущающихся сеток $\{T_N, T_{2N}, \dots, T_{2^n N}\}$ и рассмотрим поведение погрешности между решениями на соседних сетках в совпадающих узлах $\Delta_i = |E(T_{2^i N}) - E(T_{2^{i-1} N})|$ при увеличении $i = 1, \dots, n$.

Рис. 6. Отражённый импульс от фотонного кристалла ФК.

Рис. 7. Зависимость погрешности на паре последовательно сгущённых сеток от порядка сгущения первой из них.

На графике рис. 7 представлена зависимость погрешности от порядка сгущения пары сеток. Аналогичный график со степенной сходимостью представлял бы из себя прямую под некоторым наклоном. Очевидно, что функция представленная на графике рис. 7 убывает быстрее, чем линейная. Поэтому в конечном итоге требуемое число узлов определяется необходимостью детально разрешить быстрые осцилляции решения внутри отрезка интегрирования (см. рис. 6). Точность $\Delta < 10^{-3}$, заведомо достаточная для практических приложений, обеспечивается уже на сетке с ~ 2000 шагов.

Для дальнейших вычислений перейдём к интенсивностям по формулам (5).

Рис. 8. Зависимость интенсивностей начального (красная пунктирная кривая) и отражённого (синяя кривая) импульсов от времени в определённой точке пространства.

На рис. 8 показаны интенсивности падающего и отраженного импульсов в фиксированной точке пространства в зависимости от времени. Масштаб графика полулогарифмический. Поэтому исходному гауссовому профилю соответствует парабола. Отраженный импульс имеет параболическую «верхушку» и линейно убывающий «хвост». Последний соответствует излучению поверхностной волны.

Определим время жизни поверхностной волны, пользуясь полученными результатами. Для этого вычислим кросс-корреляционную функцию. В эксперименте непосредственно измеряется кросс-корреляционная функция падающего и отраженного импульсов. Она имеет следующий вид:

$$C(t) = \int_{-\infty}^{+\infty} I_r(t)I(t-\tau)dt, \quad (8)$$

где $I_r(t)$ и $I(t-\tau)$ — интенсивности падающего и отраженного импульсов в определенной точке пространства моменты времени t и $t-\tau$ соответственно.

Если поверхностная волна не формируется, то отраженный импульс имеет гауссов профиль («хвост» отсутствует). Тогда свертка (8) также является гауссовой функцией.

Если отраженный импульс имеет экспоненциально затухающий «хвост», то аналогичный «хвост» имеет и кросс-корреляционная функция. В полулогарифмическом масштабе этот «хвост» превращается в прямую линию. По ее наклону можно вычислить время жизни поверхностной волны.

Опишем процедуру численного расчета кросс-корреляционной функции. Будем использовать формулу трапеций. Зададим начальный импульс на расширенном временном интервале, при этом сохраняя шаг сетки по времени. После чего вычислим значение кросс-корреляционной функции для каждого сдвига отраженного импульса относительно начального

$$C(\tau_m) = \sum_{n=1}^N [(I_r)_n I_{n+m-1} + (I_r)_{n+1} I_{n+m}] \frac{\tau}{2}, \quad (9)$$

где $1 \leq m \leq 2N+1$.

В итоге получаем значения кросс-корреляционной функции на временной сетке $T_m = \{\tau_m, m = 1, \dots, 2N+1\}$.

Расчетная кросс-корреляционная функция представлена на рис. 9. Рассмотрим затухающий хвост кросс-корреляционной функции. На логарифмической шкале он представляет из себя прямую, что подтверждает его экспоненциальное затухание.

Рис. 9. Кросс корреляционная функция для ФК (синяя линия), для двух начальных импульсов (желтая линия) и затухающая экспонента (красная линия).

Для вычисления времени жизни поверхностной волны найдём наклон данного хвоста. Расчет дает значение $\tau = 832$ фс. Такое время жизни достаточно велико, поэтому данная структура представляется перспективной.

Заключение

В работе предложен и реализован новый метод решения нестационарных задач фотоники о прохождении излучения через фотонный кристалл. Данный метод позволяет учитывать наклонное падение, а также закон дисперсии и форму импульса, заданные в произвольном виде.

Данный метод имеет сверхбыструю сходимость, что позволяет получить решение с высокой точностью даже на умеренных сетках с небольшими числами шагов.

Проведена верификация предложенного метода сравнением с точным решением для импульса, отраженного от границы раздела сред. Подтверждена сверхстепенная сходимость данного метода.

Проведены прикладные расчеты для конкретного фотонного кристалла. Погрешность определения отраженного импульса и сходимость нестационарного матричного метода для данной задачи подтверждена аналогом метода Ричардсона.

Отражённый импульс демонстрирует возбуждение волны Блоха в данной структуре. Найдено время жизни поверхностной волны. Оно разумно согласуется с оценкой по добротности резонанса в спектре отражения.

Статья подготовлена при поддержке РФФИ (проект 19-01-00593А).

Список литературы

1. *Solymar, L., Shamonina, E.* Waves in Metamaterials. N.Y.: Oxford University Press, 2000.
2. *Engheta, N., Ziolkowski, R.W.* Metamaterials: Physics and Engineering Explorations / Wiley-IEEE Press, 2006.
3. *Домбровская Ж. О., Мухартова Ю. В., Боголюбов А. Н., Боголюбов Н. А.* Математическое моделирование электродинамических систем на основе метаматериалов // Физические основы приборостроения. 2019. Т. 8. № 3. С. 7–18.
4. *Yee, K.S.* Numerical Solution of Initial Boundary Value Problems Involving Maxwell's Equations in Isotropic Media // IEEE Transactions of Antennas and Propagations. 1966. Vol. 14. No. 3. P. 302–307
5. *Taflove, A., Johnson, S.G., and Oskooi, A.* Advances in FDTD Computational Electromagnetics: Photonics and Nanotechnology / Artech House, London, UK, Norwood, Massachusetts, USA 2013
6. *Dombrovskaya, Zh. O., Belov, A. A.* Difficulties Faced by Yee's Scheme in Photonics Problems // Journal of Physics: Conference Series. 2020. Vol. 1461. P. 012032.
7. *Berberman, D. W.* Optics in Stratified and Anisotropic Media: 4×4-Matrix Formulation // Journal of the Optical Society of America. 1972. Vol. 62. No. 4. P. 502.
8. *Белов А. А., Домбровская Ж. О.* Бикомпактная разностная схема для уравнений Максвелла в слоистых средах // Доклады российской академии наук. 2020. Т. 492. № 1. С. 15–19. DOI: 10.31857/S2686954320020034
9. *Trefethen, L.N., Weideman, J.A.C.* The Exponentially Convergent Trapezoidal Rule // SIAM Review. 2014. Vol. 56. No. 3. P. 385–458
10. *Калиткин Н. Н., Колганов С. А.* Квадратурные формулы с экспоненциальной сходимостью и функции Ферми-Дирака // Доклады российской академии наук. 2017. Т. 473. № 4. С. 401–403.
11. *Калиткин Н. Н., Белов А. А.* Аналог метода Рундсона для логарифмически сходящегося счета на установление // Доклады Академии наук. 2013. Т. 452. № 3. С. 261–265
12. *Белов А. А., Калиткин Н. Н.* Экономичные методы численного интегрирования задачи Коши для жестких систем ОДУ // Дифференциальные уравнения. 2019. Т. 55. № 7. С. 907–918.
13. *Popkova, A. A., Chezhegov, A. A., Soboleva, I. V. et al.* Difficulties Faced by Yee's Scheme in Photonics Problems // Journal of Physics: Conference Series. 2020. Vol. 1461. P. 012134.

Поступила 27 мая 2020 г.

SUPERFAST METHOD FOR ONE-DIMENSIONAL PHOTONIC PROBLEMS COMPUTATION

A. A. BELOV, A. N. BOGOLYUBOV, ZH. O. DOMBROVSKAYA, AND S. O. ZHBANNIKOV

doi: 10.25210/jfop-2002-002009

A new method is proposed and implemented for solving unsteady problems of photonics on the passage of radiation through a photonic crystal. This method allows you to take into account the inclined incidence, as well as the law of dispersion and the shape of the pulse, given in an arbitrary form.

This method has ultrafast convergence, which allows us to obtain a solution with high accuracy even on moderate grids with a small number of steps. The proposed method has been verified by comparison with the exact solution for the pulse reflected from the interface. The super-degree convergence of this method is confirmed. The error of the solution on a grid of 60 nodes is 10^{-3} .

Calculations were carried out for a specific photonic crystal, which is a structure including 7 SiO_2 and Ta_2O_5 bilayers 160 and 112 nm thick, respectively. On top of the last Ta_2O_5 layer is a SiO_2 coating 260 nm thick. The reflected pulse on a grid of 2000 nodes was found with a relative error. The error in determining the reflected pulse and the convergence of the non-stationary matrix method for this problem is confirmed by the analog of the Richardson method.

The reflected pulse demonstrates the excitation of the Bloch wave in this structure. The lifetime of the detected wave was found by numerically calculating the cross-correlation function for the reflected pulse and is equal to a femtosecond. The results are in good agreement with the estimate through the half-width of the dip in the PC spectrum in the excitation of the surface wave.

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ФИЗИЧЕСКИХ ПРОЦЕССОВ

УДК 621.391.01

ПОМЕХОУСТОЙЧИВОСТЬ ОПТИМАЛЬНОГО ПОСИМВОЛЬНОГО ПРИЕМА СИГНАЛОВ В НЕДВОИЧНЫХ ПОЛЯХ ГАЛУА

© Автор, 2020

doi: 10.25210/jfor-2002-010015

Назаров Л. Е. — гл.н.с., д.ф.-м.н., Фрязинский филиал Института радиотехники и электроники им. В. А. Котельникова РАН, Фрязино. E-mail: nazarov@sunclass.ire.rssi.ru

Аннотация

Приведено описание алгоритма оптимального посимвольного приема сигнальных конструкций на основе блочных помехоустойчивых кодов в недвоичных полях Галуа $GF(2^m)$. Показано, что основу разработанного алгоритма посимвольного приема составляет спектральное преобразование в базисе Уолша-Адамара. Результирующая сложность разработанного алгоритма посимвольного приема определяется размерностью дуального кода, что обуславливает перспективность его применения для блочных помехоустойчивых кодов с низкой избыточностью. Произведено исследование вероятностных характеристик рассматриваемого алгоритма посимвольного приема путем его моделирования для сигнальных конструкций на основе ортогональных сигналов и кодов с проверкой на четность.

Ключевые слова: помехоустойчивость, поля Галуа, сигналы, посимвольный прием, помехоустойчивые коды

Abstract

The focus of this paper is directed towards the development and investigation of the characteristics of optimal symbol-by-symbol decoding algorithms for signal constructions based on signals and on error-correcting codes in non-binary Galua fields $GF(2^m)$. The base of the developed symbol-by-symbol decoding algorithm is Fast Hadamard Transformation with dimension of Galua fields m . The complexity of decoding algorithm is determined by dimension of dual codes in non-binary fields. The computer simulations for developed symbol-by-symbol decoding for investigated signal constructions and for Additive White Gaussian Noise are performed.

Keywords: noise-immunity, non-binary Galua fields, signals, symbol-by-symbol decoding, error-correcting codes, Fast Hadamard Transformation

Введение

Помехоустойчивые коды (коды, корректирующие ошибки) используются для повышения надежности передачи информации по проводным и беспроводным каналам с помехами [1–3].

Известны два метода оптимального приема сигнальных конструкций, соответствующих помехоустойчивым кодам — метод посимвольного приема, минимизирующий вероятность ошибки на информационный бит P_6 , и метод максимального правдоподобия, минимизирующий вероятность ошибки $P_{ош}$ на кодовое слово [2,3].

Известен ряд алгоритмов посимвольного приема сигнальных конструкций, соответствующих линейным помехоустойчивым кодам в двоичном поле $GF(2)$, например, на основе использования решетчатой структуры кодов, а также на основе спектрального преобразования в базисе Уолша-Адамара [2–4]. Эти алгоритмы используются при итеративном приеме наиболее эффективных относительно вероятностных характеристик кодов — турбо-кодов, низкоплотностных кодов [3].

Актуальной является проблема разработки и исследования алгоритмов посимвольного приема сигналов в недвоичных полях $GF(2^m)$. Этот подход позволяет расширить класс кодовых конструкций, что согласуется с направлением развития теории класса помехоустойчивых кодов в недвоичных полях Галуа, включающих коды Рида-Соломона и низкоплотностные коды, интенсивно используемые в приложениях [3,5–8].

1. Постановка задачи

Пусть $H = (h_{ij}; 0 \leq i \leq k-1, 0 \leq j \leq n-1)$ — проверочная матрица блочного кода, $\vec{B} = (b_j; 0 \leq j \leq n-1)$ — кодовое слово с кодовыми символами поля $GF(2^m)$.

Кодовое слово $\vec{B}$ является решением уравнения в поле $GF(2^m)$ [1]

$$H\vec{B}^T = 0. \quad (1)$$

Элементы b_j поля представляются многочленами [1]

$$b_j(x) = \sum_{p=0}^{m-1} a_p(a_j) x^p, a_p(a_j) \in GF(2). \quad (2)$$

Кодовые символы $b_j, j = 0, 1, \dots, n$ сопоставляются сигналам, которые передаются по физическим каналам. На вход приемного устройства поступает реализация $\vec{Y} = (\dot{y}_l; 0 \leq l \leq n-1)$, где $\dot{y}_l$ — «мягкие» (многоуровневые) решения (в общем случае комплексные) с демодулятора сигналов.

Оптимальное посимвольное правило приема заключается в вычислении апостериорных вероятностей $x_\beta(j) = \Pr(b_j = \beta | \vec{Y})$, $\beta \in GF(2^m)$ и коэффициентов $a_p(b_j) \in GF(2)$, $p = 0, 1, \dots, m-1$ [3].

С использованием $x_\beta(j)$ принимаются решения относительно оценки символов $\hat{b}_j$ и значений коэффициентов $\hat{a}_p(b_j)$

$$\hat{a}_j = \max_{\beta \in GF(2^m)} (x_\beta(j)). \quad (3)$$

Апостериорные вероятности кодовых символов b_j задаются выражением [3]

$$\Pr(b_j = \beta | \vec{Y}) = \sum_{\vec{B}: b_j = \beta} \Pr(\vec{B} | \vec{Y}) = \sum_{\vec{B}: b_j = \beta} \Pr(\vec{B}) p(\vec{Y} | \vec{B}) / p(\vec{Y}). \quad (4)$$

Здесь $\Pr(\vec{B} | \vec{Y})$ — условная вероятность кодового слова $\vec{B}$ для $\vec{Y}$.

Функция правдоподобия $p(\vec{Y} | \vec{B})$ в (4) определяется моделью физического канала, для канала без памяти $p(\vec{Y} | \vec{B}) = \prod_{j=0}^n p(\dot{y}_j | b_j)$ [2]. Априорные вероятности сообщений полагаются равными $\Pr(\vec{B}) = 1 / 2^{mk}$.

Сложность реализации (4), определяемая требуемым объемом вычислительных операций, оценивается соотношением $P_1 \approx 2^{mk}$ и представляет трудноразрешимую проблему [2].

Суть задачи — дать описание разработанного эффективного относительно сложности реализации алгоритма оптимального посимвольного приема для сигнальных конструкций, соответствующих помехоустойчивым кодам в двоичных полях $GF(2^m)$, а также привести результаты моделирования разработанного алгоритма приема с целью оценки вероятностных характеристик для ряда сигнальных конструкций, соответствующих кодам с проверкой на четность.

2. Алгоритм оптимального посимвольного приема сигналов

Введем в рассмотрение множество функций [10, 11]

$$\omega_b(a) = \exp(j\pi(a \times b) \cdot d). \quad (5)$$

Здесь a, b, d — элементы поля $GF(2^m)$; $a \times b = [a(x)b(x)]$ — произведение элементов a, b в поле $GF(2^m)$, задаваемое по модулю неприводимого многочлена $\gamma(x)$ степени m ; $a \cdot d = \sum_{i=0}^{m-1} a_i \cdot d_i$.

Функции $\omega_b(a)$ принимают значения ± 1 . Полагаем, что элементы b, a определяют номер и аргумент функции $\omega_b(a)$.

При фиксированном $b \neq 0$ и для $a_i, i = 0, 1, \dots, 2^m - 1$ произведения $a_i \times b$ принимают все значения поля $GF(2^m)$ и функции $\omega_b(a)$ эквивалентны функциям Уолша $W_b(a)$ с перемешиванием их номеров и аргументов. Функции Уолша с периодом 2^m имеют вид $W_b(a) = \exp(j\pi a \cdot b) = \exp(j\pi \sum_{i=0}^{m-1} a_i b_i)$, где a_i, b_i — двоичное представление чисел a и b . Справедливы условия ортогональности [10]

$$\sum_{a \in GF(2^m)} W_b(a) = \begin{cases} 0, & b \neq 0, \\ 2^m, & b = 0, \end{cases} \quad (6)$$

$$\sum_{a \in GF(2^m)} W_b(a) W_c(a) = \sum_{a \in GF(2^m)} W_{b+c}(a) = \begin{cases} 0, & b \neq c, \\ 2^m, & b = c. \end{cases} \quad (7)$$

Функции $\omega_b(a)$ также ортогональны в области их определения, их полное множество представляет ортогональный базис, по которому можно разложить дискретные функции длительностью 2^m в ряд [10].

Для базисной системы функций Уолша разработаны производительные алгоритмы быстрого спектрального преобразования, которые также могут быть применены к перемеженной базисной системе функций (5).

Пусть неприводимый многочлен, порождающий поле $GF(2^m)$, имеет вид $\gamma(x) = 1 + x^k + x^m$ ($1 \leq k \leq m-1$), элемент поля d имеет компоненту на $(k-1)$ -ой позиции $d_{k-1} = 1$ и нулевые на позициях $i \neq k-1$. В этом случае перемещение номеров b, b' для функций $W_b(a) = w_b(a)$ определяется соотношением для компонент b_i, b'_i , $i = 0, 1, \dots, m-1$ [10,11]

$$b'_i = \begin{cases} b_{k-i-1}, & 0 \leq i \leq k-1, \\ b_{m+k-i-1}, & k \leq i \leq m-1. \end{cases} \quad (8)$$

Алгоритм оптимального посимвольного приема сигналов в поле $GF(2^m)$ включает три этапа [10,11].

На первом этапе выполняется спектральное преобразование в базисе $\omega_b(a)$ с размерностью 2^m над последовательностью «мягких» решений $p(y_i | b(l))$ (используется алгоритм быстрого спектрального преобразования в базисе Уолша-Адамара с учетом перемещения спектральных компонент (8))

$$C_l(r) = \sum_{b(l) \in GF(2^m)} p(y_i | b(l)) w_{b(l)}(r). \quad (9)$$

Здесь $l = 0, 1, \dots, n-1$ — номер позиции кодовых символов, $r \in GF(2^m)$.

Верным является обратное преобразование $p(y_i | b(l)) = \sum_{r \in GF(2^m)} C_l(r) w_{b(l)}(r) / 2^m$.

На втором этапе вычисляется спектральное множество $\{T_l(\lambda)\}$

$$T_l(\lambda) = \sum_{\beta \in GF(2^m)} \Pr(b(l) = \beta | \vec{Y}) w_{\beta}(\lambda), \quad \lambda \in GF(2^m). \quad (10)$$

Множество $\{T_l(\lambda)\}$ вычисляется с использованием величин $C_l(r)$ (9) и множества кодовых слов R дуального кода C_H с параметрами $(n, n-k)$ [11]

$$T_l(\lambda) = \left(\sum_{r_p \in C_H} \prod_{p=0}^{n-1} C_p(r_p) C_l(r_l - \lambda) / C_l(r_l) \right) / \left(\sum_{r_p \in C_H} \prod_{p=0}^{n-1} C_p(r_p) \right). \quad (11)$$

Сложность вычисления соотношения (10) оценивается приведенным значением P_1 , сложность вычисления (11) оценивается как $P_2 \cong 2^{m(n-k)}$, для значений $n-k \ll k$ справедливо условие $P_2 \ll P_1$.

Доказательство соотношения (11) основано на использовании свойств ортогональности (6), (7) для функций $\omega_b(a)$ и приведено в [10,11].

На третьем этапе итерации вычисляются апостериорные вероятности $\Pr(b(l) = \beta | \vec{Y})$ с использованием обратного спектрального преобразования над $T_l(\lambda)$ (12) (используется алгоритм быстрого спектрального преобразования Уолша-Адамара с учетом перемещения (8))

$$\Pr(b(l) = \beta | \vec{Y}) = \sum_{\lambda \in GF(2^m)} T_l(\lambda) \omega_{\beta}(\lambda). \quad (12)$$

Наиболее простым является алгоритм посимвольного приема для кодов с проверкой на четность. В этом случае объем дуального кода C_H с параметрами $(k+1, 1)$ равен 2^m , его кодовые слова — последовательности одинаковых элементов $a \in GF(2^m)$ длительностью $k+1$ [8,12].

3. Результаты моделирования

Ниже приведены вероятностные характеристики (вероятности ошибки P_e) в зависимости от отношения сигнал/помеха E_b / N_0 для канала с аддитивным белым гауссовским шумом при приеме сигнальных конструкций на основе ортогональных сигналов и кодов с проверкой на четность в полях $GF(2^3)$, $GF(2^4)$ и $GF(2^6)$, порождающие многочлены которых приведены в табл. 1 [1]. Здесь E_b — энергия сигналов на бит, N_0 — спектральная плотность (односторонняя) шума.

Таблица 1. Порождающие многочлены $\gamma(x)$ для полей $GF(2^m)$

Поле	$GF(2^3)$	$GF(2^4)$	$GF(2^6)$
$\gamma(x)$	$x^3 + x + 1$	$x^4 + x + 1$	$x^6 + x + 1$

Вероятностные характеристики получены путем моделирования приведенного алгоритма оптимального посимвольного приема при передаче кодовых слов с информационным объемом ≈ 100 и ≈ 50 битов.

При выполнении моделирования производится интервальная оценка вероятности P_e , вычисляя частоту $w = x/u$. Здесь x — число ошибочных решений в последовательности переданных битов u . Требуемое значение u определялось размером доверительного интервала $[0.5P_e, 1.5P_e]$, вероятностью P_e , доверительной вероятностью $P_{\text{доп}} = 0.95$ [12].

Для класса ортогональных сигналов объемом $M = 2^k$ известно аналитическое выражение относительно P_e [2]

$$P_e = M \int_{-\infty}^{\infty} \left[F\left(x - \sqrt{2E/N_0}\right) F^{M-2}(x) \exp(-x^2/2) / \sqrt{8\pi} \right] dx. \quad (13)$$

Здесь $E = E_b k$ — энергия сигналов в канале, $F(x)$ — интеграл ошибок [2].

Кривые 1 на рис. 1-3 соответствуют вероятностям P_e для ансамбля ортогональных сигналов объемом $M = 8$, $M = 16$ и $M = 64$, вычисленным с использованием соотношения (13). При увеличении объема M вероятности ошибки P_e монотонно уменьшаются для фиксированных значений E_b/N_0 .

На рис. 1 приведены вероятностные кривые, полученные в результате моделирования алгоритма посимвольного приема (9)-(12) сигнальных конструкций на основе ортогональных сигналов ($M = 8$) и кода с проверкой на четность в поле $GF(2^3)$ (кривая 2, $k = 17$ — информационный объем кода 51 бит; кривая 3, $k = 33$ — информационный объем кода 99 бит). Видно, что применение алгоритма посимвольного приема для этой сигнальной конструкции обуславливает энергетический выигрыш по отношению к кривой 1 — для $P_e = 10^{-5}$ выигрыш достигает 1.9 дБ.

Рис. 1. Вероятности ошибки P_e при посимвольном приеме: 1 — ансамбль ортогональных сигналов ($M = 8$); 2 — сигнальная конструкция на основе кода в поле $GF(2^3)$ (информационный объем кода 51 бит); 3 — сигнальная конструкция на основе кода в поле $GF(2^3)$ (информационный объем кода 99 битов).

Рис. 2. Вероятности ошибки P_e при посимвольном приеме: 1 — ансамбль ортогональных сигналов ($M = 16$); 2 — сигнальная конструкция на основе кода в поле $GF(2^4)$ (информационный объем кода 48 бит); 3 — сигнальная конструкция на основе кода в поле $GF(2^4)$ (информационный объем кода 100 битов).

Рис. 3. Вероятности ошибки P_6 при посимвольном приеме: 1 — ортогональные сигналы ($M = 64$); 2 — сигнальная конструкция на основе кода в поле $GF(2^6)$ (информационный объем кода 48 бит); 3 — сигнальная конструкция на основе кода в поле $GF(2^6)$ (информационный объем кода 102 бит).

На рис. 2 даны вероятностные кривые для алгоритма посимвольного приема сигнальных конструкций на основе ортогональных сигналов ($M = 16$) и кода с проверкой на четность в поле $GF(2^4)$ (кривая 2, $k = 12$ — информационный объем кода 48 битов; кривая 3, $k = 17$ — информационный объем кода 100 битов). В этом случае применение алгоритма посимвольного приема обуславливает энергетический выигрыш по отношению к кривой 1, достигающей 1.5 дБ для $P_6 = 10^{-5}$.

На рис. 3 даны вероятностные кривые для алгоритма оптимального посимвольного приема сигнальных конструкций на основе ортогональных сигналов ($M = 64$) и кода с проверкой на четность в поле $GF(2^6)$ (кривая 2, $k = 8$ — информационный объем кода 48 битов; кривая 3, $k = 17$ — информационный объем кода 102 бита). Применение алгоритма посимвольного приема для этой сигнальной конструкции обуславливает энергетический выигрыш по отношению к кривой 1 до 1.7 дБ при $P_6 = 10^{-5}$.

Заключение

Приведено описание алгоритма оптимального посимвольного приема сигнальных конструкций на основе блоковых помехоустойчивых кодов в недвоичных полях $GF(2^m)$. Основу алгоритма посимвольного приема составляет спектральное преобразование в базисе Уолша-Адамара, размерность которого определяется размерностью поля 2^m . Сложность алгоритма посимвольного приема определяется размерностью дуального кода, что обуславливает перспективность его применения для блоковых помехоустойчивых кодов с высокой кодовой скоростью.

Исследование вероятностных характеристик при использовании алгоритма посимвольного приема произведено путем его моделирования для сигнальных конструкций на основе кодов с проверкой на четность в полях $GF(2^3)$, $GF(2^4)$ и $GF(2^6)$. При увеличении размерности поля m повышается помехоустойчивость передачи информации по каналу с аддитивным белым гауссовским шумом, в частности, для вероятности ошибки $P_6 = 10^{-5}$ энергетический выигрыш достигает 1.5...1.9 дБ по отношению к ортогональным сигналам без рассматриваемого кодирования. Это обуславливает перспективность применения исследуемого алгоритма посимвольного приема в приложениях.

Работа выполнена при поддержке РФФИ (проект № 20-07-00525).

Список литературы

1. Питерсон У., Уэлдон Э. Коды, исправляющие ошибки. М.: Мир. 1976. 594 с.
2. Зюко А.Г., Фалько А.И., Панфилов И.П., Банкет В.Л., Иващенко П.В. Помехоустойчивость и эффективность систем передачи информации. М.: Радио и связь, 1985. 272 с.
3. Li, J., Lin, S., Abdel-Chaffar, K., Ryan, W.E., and Costello, D.J. Jr. LDPC Code Designs, Constructions, and Unification. Cambridge. University Press. United Kingdom. 2017. P. 248. ISBN: 978-1-107-17568-6
4. Смольянинов В.М., Назаров Л.Е. Применение спектрального преобразования в базисе Уолша при оптимальном посимвольном приеме сигналов, основанных на линейных кодах // Радиотехника и электроника. 1997. Т. 42. № 10. С. 1214–1219.
5. Steiner, F., Bocherer, G., and Liva, G. Bit-Metric Decoding of Non-Binary LDPC Codes with Probabilistic Amplitude Shaping // IEEE Communications Letters. 2018. Vol. 22. No. 11. P.2210–2213. DOI:10.1109/LCOMM.2018.2870180

6. Yeo, S., Park, I.-C. Improved Hard-Reliability Based Majority-Logic Decoding for Non-Binary LDPC Codes // IEEE Transactions on Information Theory. 2018. Vol. 64. No. 7. P. 5170–5178. DOI: 10.1109/LCOMM.2016.2623783
7. Kaipa, K. An Improvement of the Asymptotic Elias Bound for Non-Binary Codes // IEEE Communications Letters. 2018. Vol. 22. No. 11. P. 2210–2213. DOI: 10.1109/TIT.2018.2806968
8. Назаров Л. Е., Шишкин П. В. Алгоритмы посимвольного приема сигналов на основе кодов с проверкой в поле $GF(2^m)$ // Журнал радиоэлектроники. 2018. № 12. Режим доступа <http://jre.cplire.ru/jre/dec18/10/text.pdf>.
9. Смольянинов В. М., Назаров Л. Е. Оптимальный посимвольный прием сигналов, основанных на линейных кодах в полях $GF(2^m)$ // Радиотехника и электроника. 1999. Т. 44. № 7. С. 838–841.
10. Назаров Л. Е., Шишкин П. В. Исследование помехоустойчивости алгоритма оптимального посимвольного приема сигналов, соответствующих кодам с проверкой на четность в не двоичных полях // Радиотехника и электроника. 2019. Т. 64. № 9. С. 910–915. DOI: 10.1134/S0033849419080138
11. Ping, Li, Chan, S., and Yeng, K.L. Efficient Soft-in-Soft-Out Sub-Optimal Decoding Rule for Single Parity Check Codes // Electronic Letters. 1997. Vol. 33. No. 19. P. 1614–1616. DOI: 10.1049/el:19971092
12. Дунин-Барковский И. В., Смирнов Н. В. Теория вероятностей и математическая статистика в технике. М.: Гостехтеориздат, 1955. 556 с.

Поступила 4 июня 2020 г.

THE NOISE-IMMUNITY OF OPTIMAL SYMBOL-BY-SYMBOL DECODING ALGORITHM IN NON-BINARY GALUA FIELDS

L. E. NAZAROV

doi: 10.25210/jfop-2002-010015

The focus of this paper is directed towards the development and investigation of the characteristics of optimal symbol-by-symbol decoding algorithms for signal constructions based on signals and on error-correcting codes in non-binary Galua fields $GF(2^m)$. The non-binary fields are constructed on the base of module primitive polynome arithmetic. The base of the developed symbol-by-symbol decoding algorithm is Fast Hadamard Transformation with dimension of Galua fields m . The complexity of decoding algorithm is determined by dimension of dual codes in non-binary fields. The decisions evaluated with application of these decoding algorithms are equivalent to the output soft decisions. The coherent algorithms for evaluation input soft symbol decisions are developed for partial signals that equivalent orthogonal Hadamard signals. The computer simulations for developed symbol-by-symbol decoding for investigated signal constructions with information volumes 50 and 100 bits for fields $GF(2^3)$, $GF(2^4)$, $GF(2^6)$ and for Additive White Gaussian Noise are performed. It shown the error-performance degradations without code concerning with single-parity-check code are about 1.5-1.9 dB for bit-error 0.00001.

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ФИЗИЧЕСКИХ ПРОЦЕССОВ

УДК 537.8, 517.958, 537.622.4

НАЧАЛЬНО-КРАЕВАЯ ЗАДАЧА ЭКРАНИРОВАНИЯ ИМПУЛЬСНЫХ ЭЛЕКТРОМАГНИТНЫХ ПОЛЕЙ ЭКРАНАМИ ИЗ ПЕРМАЛЛОЯ ДЛЯ ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ

© Авторы, 2020

doi: 10.25210/jfop-2002-016027

Ерофеенко В. Т. — д.ф.-м.н., проф., гл.н.с., научно-исследовательская лаборатория математических методов защиты информации, Учреждение Белорусского государственного университета «Научно-исследовательский институт прикладных проблем математики и информатики», Беларусь, Минск. E-mail: bsu_erofeenko@tut.by

Урбанович А. И. — к.ф.-м.н., доцент факультета прикладной математики и информатики, Учреждение образования «Белорусский государственный университет», Беларусь, Минск, E-mail: urbanovich@gmail.com

Аннотация

Разработана математическая модель, которая описывает проникновение импульсных электромагнитных полей через плоский экран, выполненный из пермаллоя. Сформулирована трёхобластная краевая задача экранирования для областей: перед экраном, за экраном и в слое экрана. Для моделирования электромагнитного поля и поля намагниченности в пермаллое используются уравнения Максвелла с нелинейным нестационарным уравнением для поля намагниченности. На плоскостях раздела сред пермаллой-вакуум при постановке задачи учитываются классические граничные условия непрерывности тангенциальных составляющих полей. Поле намагниченности на плоскостях экрана полагается равным нулю. Трёхобластная нелинейная задача преобразована в однообластную краевую задачу для экрана. Граничные условия сопряжения на плоскостях экрана преобразованы в односторонние граничные условия. С помощью нормирований введены физически безразмерные величины и функции, характеризующие экран и электромагнитные поля. Для численной реализации однообластная краевая задача преобразована в физически безразмерную начально-краевую нелинейную задачу с соответствующими граничными и начальными условиями. Приведена формула коэффициента эффективности экранирования для импульсных полей.

Ключевые слова: плоский экран, пермаллой, импульсные электромагнитные поля, нелинейная начально-краевая задача, экранирование, математические модели, эффективность экранирования

Abstract

A mathematical model has been developed to describe the penetration of pulsed electromagnetic fields through half at permalloy shield. The boundary value problem of screening has been formulated for three domains: in front of the shield, behind the shield, and within a layer of the shield. To simulate an electromagnetic field and a magnetization field in permalloy, Maxwell equations have been used together with a nonlinear stationary equation for the magnetization field. At the interface of the two media permalloy-vacuum, the classical boundary conditions for continuity of the tangential field components have been considered. A field of magnetization at the shield planes is zero. The three domain on linear problem is transformed to the single-domain boundary-value problem for the shield. The boundary conditions for interface in gap the shield planes have been transformed to the one-sided boundary conditions. By normalizations, the physically dimensionless quantities and functions have been introduced to characterize the shield and the electromagnetic fields. For numerical realization, the single-domain boundary-value problem has been transformed to the physically dimensionless initial boundary-value problem with the corresponding boundary and initial conditions. The formula for the screening efficiency of pulsed fields has been given.

Keywords: flat shield, permalloy, pulsed electromagnetic fields, nonlinear initial boundary-value problem, screening, mathematical models, screening efficiency

Введение

Разработка математических методов моделирования проникновения электромагнитных волн через экраны является актуальным направлением современной радиационной физики [1–3]. В настоящее время для решения проблем экранирования электромагнитного излучения композитными экранами используются как однородные вещества, так и нанокompозитные структуры, составленные из различных компонентов. Следует отметить быстрорастущее использование углеродных наноструктур, например, многостенных углеродных нанотрубок с добавками из различных материалов. Но в публикациях, в основном, проводятся экспериментальные исследования. Можно выделить работу [4], где исследовались эффекты пропу-

скания, отражения и поглощения микроволнового электромагнитного излучения композитными материалами, содержащими добавки из магнитных материалов. Авторы этой работы обратили внимание на изменение эффективности экранирования за счет взаимодействия ферритных магнитных моментов композитного материала с магнитной составляющей электромагнитного излучения.

Одним из веществ, используемых для создания экранов, является пермаллой [5, 6]. Это магнитомягкий, то есть, легконамагничиваемый сплав никеля с железом, обладающий высокой магнитной проницаемостью и имеющий малую коэрцитивную силу. Пермаллой легко намагничивается благодаря тому, что он является магнитоизотропным материалом. Из-за того, что в пермаллое отсутствует анизотропия, при воздействии на него внешнего поля внутри материала не возникают механические напряжения, что очень важно при его использовании в качестве наполнителя в наноструктурах, например, в нанотрубках. Пермаллой обладает очень высокими магнитными свойствами в слабых полях. Пермаллоидные материалы относят к прецизионным сплавам, то есть, к материалам с редким сочетанием физических и химических свойств. Пермаллой характеризуется полем намагниченности $\vec{M}$, которое при импульсных полях, то есть в нестационарном режиме, сильно изменяется во времени нелинейным образом. Отметим, что в работах [5, 6] изучаются краевые задачи для экранов из пермаллоя, описываемого линеаризованными дифференциальными уравнениями.

В основе задачи экранирования лежит определение электромагнитного поля, прошедшего через экран, и сравнение его с полем, воздействующим на экран. В этом случае рассчитывается эффективность экранирования, которая равна отношению амплитуды поля, падающего на экран, к амплитуде поля, прошедшего через экран. Определенный таким образом коэффициент эффективности экранирования показывает, во сколько раз ослабевает внешнее электромагнитное излучение при прохождении через плоский экран из области $D_1 (z < 0)$ в область $D_2 (z > \Delta)$, где экран определен слоем $D (0 < z < \Delta)$.

Постановка задачи экранирования предполагает рассмотрение падающего на экран электромагнитного поля, поля отраженного от него, поля внутри экрана и поля, прошедшего через экран. Это делает задачу достаточно сложной. Если же учесть импульсный характер электромагнитного поля, то параметры, характеризующие магнитные свойства среды, начинают изменяться во времени, при этом, в общем случае, задача является нелинейной, поэтому получение аналитического решения становится невозможным. Для решения краевых задач используются аналитические [7], аналитико-численные [8] и численные методы [9].

В литературе разрабатываются методы решения краевых задач прохождения электромагнитных волн через экраны из различных композитных материалов: биизотропных [10], с пространственной дисперсией [11], из пермаллоя [6, 12], из киральных [3, 13] и сверхпроводящих [14] материалов. В данной работе разработана математическая модель нелинейной краевой задачи экранирования импульсных полей плоским экраном из пермаллоя. Сформулирована трехобластная краевая задача экранирования для трех областей, а именно: первая область — вакуумное полупространство перед экраном, вторая область — плоский слой, заполненный пермаллом, третья область — вакуумное полупространство за экраном. Для формулировки задачи разработаны граничные условия сопряжения для нормальных и касательных составляющих полей на плоских поверхностях раздела сред. Исходная трехобластная нелинейная краевая задача преобразована в однообластную нелинейную начально-краевую задачу для слоя из пермаллоя. Полученная для численных расчетов начально-краевая задача представляет собой систему из шести скалярных параболических нелинейных дифференциальных уравнений по временной и пространственной переменным с граничными условиями третьего и первого рода на лицевых поверхностях экрана.

1. Нелинейные дифференциальные уравнения, описывающие распространение электромагнитных полей в материале из пермаллоя

Рассмотрим плоский слой $D (0 < z < \Delta)$, заполненный пермаллом. Для моделирования электромагнитного поля $\vec{E}(\vec{r}, t)$, $\vec{H}(\vec{r}, t)$ и поля намагниченности $\vec{M}(\vec{r}, t)$, распространяющихся в слое D , воспользуемся следующей системой дифференциальных уравнений [5]:

$$\text{rot } \vec{E} = -\mu_0 \frac{\partial}{\partial t} (\vec{H} + \vec{M}), \quad (1)$$

$$\text{rot } \vec{H} = \vec{j}, \quad (2)$$

$$\frac{\partial \vec{M}}{\partial t} = \gamma \vec{P}, \quad \vec{P} = \vec{M} \times (\vec{H} + \dot{a} \vec{M} - \dot{g} \vec{M} \times \vec{H}), \quad (3)$$

где μ_0 — магнитная постоянная; $\dot{\sigma}$ — проводимость пермаллоя; величины $\dot{\gamma}$, $\dot{a}$, $\dot{g}$ определяют масштабы

изменения намагниченности $\vec{M}(z, t)$ и её связь с напряженностью магнитного поля $\vec{H}(z, t)$; постоянные имеют физические размерности

$$[\dot{\gamma}] = \frac{M}{A \cdot c}, [\dot{a}] = M^2, [\dot{g}] = \frac{M}{A}, [\dot{\sigma}] = \frac{C_M}{M}, \mu_0 = 4\pi \cdot 10^{-7} \frac{ГН}{M},$$

где $\times$ — означает векторное произведение.

Преобразуем уравнения (1)-(3) в систему нелинейных уравнений параболического типа. Из уравнения (1) исключим вектор $\vec{E}$ с помощью уравнения (2) и временную производную от вектора намагниченности $\vec{M}$ с помощью первого уравнения (3). Получим систему уравнений, которая в дальнейшем используется для постановки краевой задачи экранирования.

$$\frac{\partial \vec{H}}{\partial t} = \frac{1}{\dot{\sigma} \mu_0} (\Delta \vec{H} - \text{grad div} \vec{H}) - \dot{\gamma} \vec{P}, \quad 0 < z < \Delta, \quad (4)$$

$$\frac{\partial \vec{M}}{\partial t} = \dot{\gamma} \vec{P}, \quad 0 < z < \Delta. \quad (5)$$

Будем искать решения системы уравнений (4), (5) в виде полей, равномерно распределённых вдоль слоя D вида

$$\vec{H} = H_x(z, t) \vec{e}_x + H_y(z, t) \vec{e}_y + H_z(z, t) \vec{e}_z, \quad \vec{M} = M_x(z, t) \vec{e}_x + M_y(z, t) \vec{e}_y + M_z(z, t) \vec{e}_z. \quad (6)$$

Найдем выражение для поля $\vec{P}(z, t)$, вычисляя следующие векторные произведения:

$$\begin{aligned} \vec{M} \times \vec{H} &= \vec{e}_x (M_y H_z - M_z H_y) + \vec{e}_y (M_z H_x - M_x H_z) + \vec{e}_z (M_x H_y - M_y H_x), \\ \vec{M} \times \Delta \vec{M} &= \vec{e}_x (M_y \Delta M_z - M_z \Delta M_y) + \vec{e}_y (M_z \Delta M_x - M_x \Delta M_z) + \vec{e}_z (M_x \Delta M_y - M_y \Delta M_x), \\ \vec{M} \times (\vec{M} \times \vec{H}) &= \vec{e}_x (M_y (M_x H_y - M_y H_x) - M_z (M_z H_x - M_x H_z)) + \\ &+ \vec{e}_y (M_z (M_y H_z - M_z H_y) - M_x (M_x H_y - M_y H_x)) + \\ &+ \vec{e}_z (M_x (M_z H_x - M_x H_z) - M_y (M_y H_z - M_z H_y)). \end{aligned}$$

С учётом соотношения

$$\Delta M_\alpha = \frac{\partial^2 M_\alpha}{\partial z^2} \quad (\alpha = x, y, z)$$

для вектора $\vec{P}$ найдем представление

$$\vec{P} = P_x(z, t) \vec{e}_x + P_y(z, t) \vec{e}_y + P_z(z, t) \vec{e}_z,$$

где

$$\begin{aligned} P_x &= \dot{a} \left(M_y \frac{\partial^2 M_z}{\partial z^2} - M_z \frac{\partial^2 M_y}{\partial z^2} \right) + M_y H_z - M_z H_y - \\ &- \dot{g} (M_y (M_x H_y - M_y H_x) - M_z (M_z H_x - M_x H_z)), \\ P_y &= \dot{a} \left(M_z \frac{\partial^2 M_x}{\partial z^2} - M_x \frac{\partial^2 M_z}{\partial z^2} \right) + M_z H_x - M_x H_z - \\ &- \dot{g} (M_z (M_y H_z - M_z H_y) - M_x (M_x H_y - M_y H_x)), \\ P_z &= \dot{a} \left(M_x \frac{\partial^2 M_y}{\partial z^2} - M_y \frac{\partial^2 M_x}{\partial z^2} \right) + M_x H_y - M_y H_x - \\ &- \dot{g} (M_x (M_z H_x - M_x H_z) - M_y (M_y H_z - M_z H_y)). \end{aligned} \quad (7)$$

Аналогичным образом вычислим слагаемые, входящие в уравнение (4)

$$\Delta \vec{H} = \frac{\partial^2 H_x}{\partial z^2} \vec{e}_x + \frac{\partial^2 H_y}{\partial z^2} \vec{e}_y + \frac{\partial^2 H_z}{\partial z^2} \vec{e}_z, \quad \text{grad div} \vec{H} = \frac{\partial^2 H_z}{\partial z^2} \vec{e}_z. \quad (8)$$

Сформулируем краевую задачу для системы уравнений (4), (5).

2. Постановка краевой задачи экранирования импульсных электромагнитных полей экраном из пермаллоя

В пространстве E^3 с электрической и магнитной постоянными ε_0, μ_0 расположен плоский экран $D(0 < z < \Delta)$ толщины Δ из пермаллоя, ограниченный плоскостями $\Gamma_1(z=0), \Gamma_2(z=\Delta)$. Из полупространства $D_1(z < 0)$ на слой D воздействует первичное электромагнитное поле $\vec{E}_0(z, t), \vec{H}_0(z, t)$. В результате в области D_1 образуется отражённое поле $\vec{E}_1(z, t), \vec{H}_1(z, t)$ и суммарное электромагнитное поле

$$\vec{E}_1(z, t) = \vec{E}_0(z, t) + \vec{E}_1(z, t),$$

$$\vec{H}_1(z, t) = \vec{H}_0(z, t) + \vec{H}_1(z, t).$$

В полупространство $D_2(z > \Delta)$ через экран D проникает поле $\vec{E}_2(z, t), \vec{H}_2(z, t)$.

Краевая задача

Для заданного первичного поля $\vec{E}_0(z, t), \vec{H}_0(z, t)$ требуется определить поля $\vec{E}_1(z, t), \vec{H}_1(z, t); \vec{E}_2(z, t), \vec{H}_2(z, t)$, соответственно в областях D_1, D_2 , и поля $\vec{E}(\vec{r}, t), \vec{H}(\vec{r}, t), \vec{M}(\vec{r}, t)$ в области D , которые удовлетворяют следующим условиям:

- уравнениям Максвелла

$$\text{rot} \vec{E}_1' = -\mu_0 \frac{\partial \vec{H}_1'}{\partial t}, \text{rot} \vec{H}_1' = \varepsilon_0 \frac{\partial \vec{E}_1'}{\partial t}, z < 0, \quad (9)$$

$$\text{rot} \vec{E}_2 = -\mu_0 \frac{\partial \vec{H}_2}{\partial t}, \text{rot} \vec{H}_2 = \varepsilon_0 \frac{\partial \vec{E}_2}{\partial t}, z > \Delta;$$

$$\frac{\partial \vec{H}}{\partial t} = \frac{1}{\sigma \mu_0} (\Delta \vec{H} - \text{grad} \text{div} \vec{H}) - \gamma \vec{P}, 0 < z < \Delta, \quad (10)$$

$$\frac{\partial \vec{M}}{\partial t} = \gamma \vec{P}, 0 < z < \Delta; \quad (11)$$

- граничным условиям сопряжения на плоскости $\Gamma_1(z=0)$

$$(\vec{E}_r - \vec{E}_{1r})|_{z=0} = 0, (\vec{H}_r - \vec{H}_{1r})|_{z=0} = 0, \quad (12)$$

$$\vec{M}|_{z=0} = 0; \quad (13)$$

- граничным условиям сопряжения на плоскости $\Gamma_2(z=\Delta)$

$$(\vec{E}_r - \vec{E}_{2r})|_{z=\Delta} = 0, (\vec{H}_r - \vec{H}_{2r})|_{z=\Delta} = 0, \quad (14)$$

$$\vec{M}|_{z=\Delta} = 0; \quad (15)$$

- условиям излучения на бесконечность в областях D_1, D_2 .

Заметим, что краевая задача (9)-(15) является трёхобластной краевой задачей для полупространств перед экраном D_1 , за экраном D_2 и для слоя D с граничными условиями сопряжения на поверхностях раздела сред Γ_1 и Γ_2 . В последующем трёхобластная краевая задача будет преобразована в однообластную краевую задачу с односторонними граничными условиями на лицевых плоскостях экрана Γ_1 и Γ_2 . При этом области D_1, D_2 исключаются из рассмотрения. Такая процедура необходима для упрощения численной реализации решения нелинейной краевой задачи (9)-(15). При постановке задачи предполагается, что векторы электрического и магнитного полей в областях D_1, D_2 параллельны экрану, т.е. поток энергии полей ортогонален экрану.

3. Импульсные электромагнитные поля, распространяющиеся в вакууме

Импульсные поля в вакууме, распространяющиеся в положительном направлении оси Oz ортогонально слою D и подчиняющиеся уравнениям Максвелла

$$\text{rot} \vec{E} = -\mu_0 \frac{\partial \vec{H}}{\partial t}, \text{rot} \vec{H} = \varepsilon_0 \frac{\partial \vec{E}}{\partial t}, \quad (16)$$

задаются формулами [14]

$$\vec{E}(z, t) = A \left(\frac{ct - z}{c\tau} \right) \vec{e}_x - B \left(\frac{ct - z}{c\tau} \right) \vec{e}_y, \quad \vec{H}(z, t) = \frac{1}{Z_0} \left(B \left(\frac{ct - z}{c\tau} \right) \vec{e}_x + A \left(\frac{ct - z}{c\tau} \right) \vec{e}_y \right), \quad (17)$$

где $A(t')$, $B(t')$ — импульсные функции, заданные на множестве $0 < t' < \infty$; $Z_0 = \sqrt{\frac{\mu_0}{\varepsilon_0}}$, τ — характерное время, характеризующее длительность импульса; c — скорость света в вакууме.

Положим $\tau = \tau_{\text{фр}}$, где $\tau_{\text{фр}}$ — время фронта импульса, время за которое значение импульса изменяется от нуля до максимального значения. Импульсные поля, распространяющиеся в отрицательном направлении оси Oz и подчиняющиеся уравнениям (16), задаются формулами

$$\vec{E}(z, t) = -A \left(\frac{ct + z}{c\tau} \right) \vec{e}_x + B \left(\frac{ct + z}{c\tau} \right) \vec{e}_y, \quad \vec{H}(z, t) = \frac{1}{Z_0} \left(B \left(\frac{ct + z}{c\tau} \right) \vec{e}_x + A \left(\frac{ct + z}{c\tau} \right) \vec{e}_y \right), \quad (18)$$

Для постановки краевой задачи проникновения электромагнитных полей через плоский экран из пермаллоя в качестве первичного поля, воздействующего на экран, рассмотрим импульсное электромагнитное поле, распространяющееся ортогонально к поверхности экрана D .

$$\vec{E}_0(z, t) = -B_0 \left(\frac{ct - z}{c\tau} \right) \vec{e}_y, \quad \vec{H}_0(z, t) = \frac{1}{Z_0} B_0 \left(\frac{ct - z}{c\tau} \right) \vec{e}_x. \quad (19)$$

В качестве импульсной функции первичного поля выберем дважды непрерывно дифференцируемую функцию $B_0(t') = Z_0 H_0 b(t')$, $\max_{0 \leq t' < 2} b(t') = 1$, где H_0 — максимальное значение напряжённости магнитного поля $\vec{H}_0$, $-\infty \leq t' \leq \infty$

$$\begin{cases} b(t') = \frac{1}{2} \sin^3 \left(\pi \left(t' - \frac{1}{2} \right) \right) + \frac{1}{2}, & 0 \leq t' \leq 2, \\ b(t') = 0 & \text{при } t' \leq 0 \text{ и } t' \geq 2. \end{cases} \quad (20)$$

Заметим, что безразмерное время фронта $t'_{\text{фр}} = 1$, а безразмерное время длительности импульса $t'_{\text{им}} = 2$. Для моделирования бесконечно дифференцируемых импульсных функций сложной конфигурации используются атомарные функции [8]. В простейшем случае можно положить

$$b(t') = \text{up}(t' - 1), \quad \text{up}(t' - 1) = 0 \quad \text{при } t' \leq 0 \text{ и } t' \geq 2.$$

В результате при воздействии поля (20) в области $D_1 (z < 0)$ образуется отражённое импульсное поле вида (18)

$$\vec{E}_1(z, t) = -A_1 \left(\frac{ct + z}{c\tau} \right) \vec{e}_x + B_1 \left(\frac{ct + z}{c\tau} \right) \vec{e}_y, \quad \vec{H}_1(z, t) = \frac{1}{Z_0} \left(B_1 \left(\frac{ct + z}{c\tau} \right) \vec{e}_x + A_1 \left(\frac{ct + z}{c\tau} \right) \vec{e}_y \right). \quad (21)$$

Суммарное поле в области D_1 будет иметь вид

$$\vec{E}_1(z, t) = \vec{E}_0(z, t) + \vec{E}_1(z, t) = -A_1 \left(\frac{ct + z}{c\tau} \right) \vec{e}_x + \left(-B_0 \left(\frac{ct - z}{c\tau} \right) + B_1 \left(\frac{ct + z}{c\tau} \right) \right) \vec{e}_y, \quad (22)$$

$$\vec{H}_1(z, t) = \vec{H}_0(z, t) + \vec{H}_1(z, t) = \frac{1}{Z_0} \left[\left(B_0 \left(\frac{ct - z}{c\tau} \right) + B_1 \left(\frac{ct + z}{c\tau} \right) \right) \vec{e}_x + A_1 \left(\frac{ct + z}{c\tau} \right) \vec{e}_y \right].$$

В область $D_2 (z > \Delta)$ проникает импульсное электромагнитное поле со структурой поля вида (17)

$$\begin{aligned} \vec{E}_2(z, t) &= A_2 \left(\frac{ct - z}{c\tau} \right) \vec{e}_x - B_2 \left(\frac{ct - z}{c\tau} \right) \vec{e}_y, \\ \vec{H}_2(z, t) &= \frac{1}{Z_0} \left(B_2 \left(\frac{ct - z}{c\tau} \right) \vec{e}_x + A_2 \left(\frac{ct - z}{c\tau} \right) \vec{e}_y \right), \end{aligned} \quad (23)$$

Импульсные функции $A_j(t')$, $B_j(t')$ полей (21), (23) определяются в результате решения краевой задачи (9)-(15).

4. Граничные условия для нормальных составляющих полей на плоскостях экрана

Рассмотрим плоскость $\Gamma(z = \text{const})$ разделяющую среды $\bar{D}_1, \bar{D}_2$, в которых распространяются поля $\vec{E}_j, \vec{H}_j, (j=1,2)$, подчиняющиеся уравнениям Максвелла

$$\text{rot} \vec{E}_j = -\frac{\partial}{\partial t} \vec{B}_j, \quad \text{rot} \vec{H}_j = \frac{\partial}{\partial t} \vec{D}_j + \vec{J}_j.$$

Граничные условия, связывающие нормальные составляющие полей на поверхности Γ , определяются формулами [7, с. 195]

$$\left((\vec{D}_2, \vec{n}) - (\vec{D}_1, \vec{n}) + \int_0^t (\vec{J}_2, \vec{n}) dt - \int_0^t (\vec{J}_1, \vec{n}) dt \right) \Big|_{\Gamma} = 0, \\ \left((\vec{B}_2, \vec{n}) - (\vec{B}_1, \vec{n}) \right) \Big|_{\Gamma} = 0, \quad (24)$$

где $\vec{n}$ — нормаль к плоскости Γ , направленная в область $\bar{D}_2$.

В соответствии с формулами (24) построим граничные условия на плоскости $\Gamma_1(z=0)$ плоского экрана D из пермаллоя, где $\bar{D}_1 = D_1, \bar{D}_2 = D$. Из уравнений (1), (2) следуют формулы

$$\vec{B}_2 = \mu_0(\vec{H} + \vec{M}), \quad \vec{D}_2 = 0, \quad \vec{J}_2 = \sigma \vec{E}.$$

В области $D_1(z < 0)$ вакуум, поэтому

$$\vec{B}_1 = \mu_0 \vec{H}_1, \quad \vec{D}_1 = \varepsilon_0 \vec{E}_1, \quad \vec{J}_1 = 0.$$

В результате граничные условия (24) на плоскости Γ_1 принимают вид

$$\left(-\varepsilon_0 (\vec{E}_1, \vec{n}) + \sigma \int_0^t (\vec{E}, \vec{n}) dt \right) \Big|_{z=0} = 0, \\ \left((\vec{H} + \vec{M}, \vec{n}) - (\vec{H}_1, \vec{n}) \right) \Big|_{z=0} = 0, \quad \vec{n} = \vec{e}_z. \quad (25)$$

Выведем граничные условия на плоскости $\Gamma_2(z=\Delta)$ экрана D , где $\bar{D}_1 = D, \bar{D}_2 = D_2$. Из уравнений (1), (2) для экрана следует $\vec{B}_1 = \mu_0(\vec{H} + \vec{M}), \vec{D}_1 = 0, \vec{J}_1 = \sigma \vec{E}$. В области $D_2(z > \Delta)$ вакуум, следует $\vec{B}_2 = \mu_0 \vec{H}_2, \vec{D}_2 = \varepsilon_0 \vec{E}_2, \vec{J}_2 = 0$.

В результате граничные условия (18) на плоскости Γ_2 примут вид

$$\left(\varepsilon_0 (\vec{E}_2, \vec{n}) - \sigma \int_0^t (\vec{E}, \vec{n}) dt \right) \Big|_{z=\Delta} = 0, \quad \left((\vec{H} + \vec{M}, \vec{n}) - (\vec{H}_2, \vec{n}) \right) \Big|_{z=\Delta} = 0, \quad \vec{n} = \vec{e}_z. \quad (26)$$

Дифференцируя интегральные уравнения (25), (26) по переменной t , получим итоговые граничные условия сопряжения на лицевых поверхностях Γ_1, Γ_2 экрана D для нормальных составляющих полей

$$\left(-\varepsilon_0 \frac{\partial}{\partial t} (\vec{E}_1, \vec{n}) + \sigma (\vec{E}, \vec{n}) \right) \Big|_{z=0} = 0, \quad \left((\vec{H} + \vec{M}, \vec{n}) - (\vec{H}_1, \vec{n}) \right) \Big|_{z=0} = 0, \\ \left(\varepsilon_0 \frac{\partial}{\partial t} (\vec{E}_2, \vec{n}) - \sigma (\vec{E}, \vec{n}) \right) \Big|_{z=\Delta} = 0, \quad \left((\vec{H} + \vec{M}, \vec{n}) - (\vec{H}_2, \vec{n}) \right) \Big|_{z=\Delta} = 0, \quad \vec{n} = \vec{e}_z. \quad (27)$$

С помощью уравнения (2) вычислим компоненты поля $\vec{E}$

$$\vec{E}(z, t) = E_x(z, t) \vec{e}_x + E_y(z, t) \vec{e}_y = \frac{1}{\sigma} \left(-\frac{\partial H_y(z, t)}{\partial z} \vec{e}_x + \frac{\partial H_x(z, t)}{\partial z} \vec{e}_y \right). \quad (28)$$

Из формул (22), (23) следует $(\vec{E}_j, \vec{n}) = 0, (\vec{H}_j, \vec{n}) = 0$, а из представления электрического поля (28) получим $(\vec{E}, \vec{n}) = 0$. Таким образом, первые условия в уравнениях (27) выполнены тождественно. В результате совокупность граничных условий (27) принимает окончательный вид

$$(\vec{H} + \vec{M}, \vec{n}) \Big|_{z=0} = 0, \quad (\vec{H} + \vec{M}, \vec{n}) \Big|_{z=\Delta} = 0, \quad \vec{n} = \vec{e}_z. \quad (29)$$

Заметим, что намагниченность среды в областях D_1, D_2 отсутствует. В связи с этим для намагниченности на плоскостях Γ_1, Γ_2 полагаются выполненными граничные условия (13), (15), откуда следуют скалярные условия

$$M_x|_{z=0}=0, M_y|_{z=0}=0, M_z|_{z=0}=0, M_x|_{z=\Delta}=0, M_y|_{z=\Delta}=0, M_z|_{z=\Delta}=0. \quad (30)$$

Из соотношений (29), (30) получим граничные условия для нормальной составляющей магнитного поля.

$$H_z|_{z=0}=0, H_z|_{z=\Delta}=0. \quad (31)$$

5. Односторонние граничные условия для касательных составляющих полей на плоскостях экрана

Преобразуем двухсторонние граничные условия сопряжения (12), (14) в односторонние граничные условия.

Для магнитных полей в областях D и D_1 на плоскости $\Gamma_1 (z=0)$ выполнены граничные условия непрерывности тангенциальных составляющих полей (12)

$$[\vec{n}, [\vec{H}, \vec{n}]]|_{z=0} = [\vec{n}, [\vec{H}_1, \vec{n}]]|_{z=0}. \quad (32)$$

Учитывая формулы (6), (22) при $z=0$ для магнитных полей в областях $D(0 \leq z \leq \Delta)$, $D_1(z \leq 0)$, преобразуем условия (32) в граничные условия для декартовых составляющих

$$H_x(0, t) = \frac{1}{Z_0} \left(B_0 \left(\frac{t}{\tau} \right) + B_1 \left(\frac{t}{\tau} \right) \right), H_y(0, t) = \frac{1}{Z_0} A_1 \left(\frac{t}{\tau} \right). \quad (33)$$

Для электрических полей в областях D и D_1 на плоскости $\Gamma_1 (z=0)$ выполнены граничные условия сопряжения (12)

$$[\vec{n}, [\vec{E}, \vec{n}]]|_{z=0} = [\vec{n}, [\vec{E}_1, \vec{n}]]|_{z=0}. \quad (34)$$

Аналогично, учитывая формулы (6), (22) при $z=0$ для электрических полей, преобразуем условия (34) в скалярные граничные условия для декартовых составляющих полей

$$E_x(0, t) = -A_1 \left(\frac{t}{\tau} \right), E_y(0, t) = -B_0 \left(\frac{t}{\tau} \right) + B_1 \left(\frac{t}{\tau} \right). \quad (35)$$

Из равенств (33), используя соотношения (35), исключим неизвестные функции A_1 и B_1 . Получим равенства

$$H_x(0, t) = \frac{1}{Z_0} \left(2B_0 \left(\frac{t}{\tau} \right) + E_y(0, t) \right), H_y(0, t) = -\frac{1}{Z_0} E_x(0, t). \quad (36)$$

С помощью формул (28) при $z=0$ исключим компоненты $E_x(0, t)$, $E_y(0, t)$ из тождеств (36). Тогда получим

$$\left(\frac{\partial H_x(z, t)}{\partial z} - Z_0 \dot{H}_x(z, t) \right) \Big|_{z=0} = -2\dot{B}_0 \left(\frac{t}{\tau} \right), \left(\frac{\partial H_y(z, t)}{\partial z} - Z_0 \dot{H}_y(z, t) \right) \Big|_{z=0} = 0. \quad (37)$$

Аналогично, для магнитных и электрических полей в областях $D(0 \leq z \leq \Delta)$, $D_2(z \geq \Delta)$ на плоскости $\Gamma_2 (z = \Delta)$ выполнены граничные условия непрерывности тангенциальных составляющих полей (14)

$$[\vec{n}, [\vec{H}, \vec{n}]]|_{z=\Delta} = [\vec{n}, [\vec{H}_2, \vec{n}]]|_{z=\Delta}, [\vec{n}, [\vec{E}, \vec{n}]]|_{z=\Delta} = [\vec{n}, [\vec{E}_2, \vec{n}]]|_{z=\Delta}. \quad (38)$$

Учитывая формулы (6), (23) при $z = \Delta$ для магнитных и электрических полей в областях D и D_2 , преобразуем условия (38) в скалярные граничные условия для декартовых составляющих

$$H_x(\Delta, t) = \frac{1}{Z_0} B_2 \left(\frac{ct - \Delta}{c\tau} \right), H_y(\Delta, t) = \frac{1}{Z_0} A_2 \left(\frac{ct - \Delta}{c\tau} \right), \quad (39)$$

$$E_x(\Delta, t) = A_2 \left(\frac{ct - \Delta}{c\tau} \right), E_y(\Delta, t) = -B_2 \left(\frac{ct - \Delta}{c\tau} \right). \quad (40)$$

С помощью формул (40) исключим функции A_2 и B_2 из равенств (39), получим

$$H_x(\Delta, t) = -\frac{1}{Z_0} E_y(\Delta, t), H_y(\Delta, t) = \frac{1}{Z_0} E_x(\Delta, t). \quad (41)$$

Учитывая соотношения (22), вычислим

$$E_x(\Delta, t) = -\frac{1}{\dot{\sigma}} \frac{\partial H_y(z, t)}{\partial z} \Big|_{z=\Delta},$$

$$E_y(\Delta, t) = \frac{1}{\dot{\sigma}} \frac{\partial H_x(z, t)}{\partial z} \Big|_{z=\Delta}.$$

После подстановки в (41) получим граничные условия на плоскости Γ_2

$$\left(\frac{\partial H_x(z, t)}{\partial z} + Z_0 \dot{\sigma} H_x(z, t) \right) \Big|_{z=\Delta} = 0, \quad \left(\frac{\partial H_y(z, t)}{\partial z} + Z_0 \dot{\sigma} H_y(z, t) \right) \Big|_{z=\Delta} = 0. \quad (42)$$

Заметим, что односторонние граничные условия (37), (42) содержат только магнитное поле, сосредоточенное в слое D .

6. Физически безразмерные параболические дифференциальные уравнения в слое из пермаллоя и граничные условия на плоскостях слоя

Введём безразмерное время $\bar{t} = \frac{t}{\tau}$ и безразмерную координату $\bar{z} = \frac{z}{\Delta}$.

Обезразмерим поля $\vec{H}(z, t)$, $\vec{M}(z, t)$, определённые в (6), полагая

$$\begin{aligned} H_x(z, t) &= H_x(\Delta \bar{z}, \tau \bar{t}) = H_0 u_1(\bar{z}, \bar{t}), \\ H_y(z, t) &= H_y(\Delta \bar{z}, \tau \bar{t}) = H_0 u_2(\bar{z}, \bar{t}), \\ H_z(z, t) &= H_z(\Delta \bar{z}, \tau \bar{t}) = H_0 u_3(\bar{z}, \bar{t}); \\ M_x(z, t) &= M_x(\Delta \bar{z}, \tau \bar{t}) = H_0 v_1(\bar{z}, \bar{t}), \\ M_y(z, t) &= M_y(\Delta \bar{z}, \tau \bar{t}) = H_0 v_2(\bar{z}, \bar{t}), \\ M_z(z, t) &= M_z(\Delta \bar{z}, \tau \bar{t}) = H_0 v_3(\bar{z}, \bar{t}). \end{aligned} \quad (43)$$

Для безразмерных функций $u_1, u_2, u_3; v_1, v_2, v_3$ построим систему безразмерных дифференциальных уравнений.

Дифференциальные уравнения

Векторное дифференциальное уравнение (11) запишем для компонент поля (6)

$$\frac{\partial M_x(z, t)}{\partial t} = \dot{\gamma} P_x(z, t), \quad \frac{\partial M_y(z, t)}{\partial t} = \dot{\gamma} P_y(z, t), \quad \frac{\partial M_z(z, t)}{\partial t} = \dot{\gamma} P_z(z, t). \quad (44)$$

Обезразмерим уравнения (44), учитывая представления (7) и (43). Получим безразмерные уравнения для пермаллоя в слое $D (0 < \bar{z} < 1)$

$$\frac{\partial v_1(\bar{z}, \bar{t})}{\partial \bar{t}} = p_1(\bar{z}, \bar{t}), \quad \frac{\partial v_2(\bar{z}, \bar{t})}{\partial \bar{t}} = p_2(\bar{z}, \bar{t}), \quad \frac{\partial v_3(\bar{z}, \bar{t})}{\partial \bar{t}} = p_3(\bar{z}, \bar{t}), \quad 0 < \bar{z} < 1, \quad 0 < \bar{t} < \infty, \quad (45)$$

где

$$\begin{aligned} p_1(\bar{z}, \bar{t}) &= F_1 \left(\frac{\partial}{\partial \bar{z}} \left(v_2 \frac{\partial v_3}{\partial \bar{z}} \right) - \frac{\partial}{\partial \bar{z}} \left(v_3 \frac{\partial v_2}{\partial \bar{z}} \right) \right) + F_2 (v_2 u_3 - v_3 u_2) + \\ &\quad + F_3 [v_2 (v_2 u_1 - v_1 u_2) - v_3 (v_3 u_1 - v_1 u_3)], \\ p_2(\bar{z}, \bar{t}) &= F_1 \left(\frac{\partial}{\partial \bar{z}} \left(v_3 \frac{\partial v_1}{\partial \bar{z}} \right) - \frac{\partial}{\partial \bar{z}} \left(v_1 \frac{\partial v_3}{\partial \bar{z}} \right) \right) + F_2 (v_3 u_1 - v_1 u_3) + \\ &\quad + F_3 [v_3 (v_3 u_2 - v_2 u_3) + v_1 (v_1 u_2 - v_2 u_1)], \\ p_3(\bar{z}, \bar{t}) &= F_1 \left(\frac{\partial}{\partial \bar{z}} \left(v_1 \frac{\partial v_2}{\partial \bar{z}} \right) - \frac{\partial}{\partial \bar{z}} \left(v_2 \frac{\partial v_1}{\partial \bar{z}} \right) \right) + F_2 (v_1 u_2 - v_2 u_1) + \\ &\quad + F_3 [v_1 (v_1 u_3 - v_3 u_1) + v_2 (v_2 u_3 - v_3 u_2)]. \end{aligned} \quad (46)$$

Здесь $F_1 = \frac{\dot{\gamma} \tau \dot{\sigma} H_0}{\Delta^2}$, $F_2 = \dot{\gamma} \tau H_0$, $F_3 = \dot{\gamma} \tau \dot{\sigma} H_0^2$ — безразмерные постоянные.

Аналогичным образом векторное дифференциальное уравнение (10) запишем покомпонентно, учитывая разложения полей (6), (8). В результате получим скалярные уравнения

$$\begin{aligned}\frac{\partial H_x(z,t)}{\partial t} &= \frac{1}{\dot{\sigma}\mu_0} \frac{\partial^2 H_x(z,t)}{\partial z^2} - \dot{\gamma} P_x(z,t), \\ \frac{\partial H_y(z,t)}{\partial t} &= \frac{1}{\dot{\sigma}\mu_0} \frac{\partial^2 H_y(z,t)}{\partial z^2} - \dot{\gamma} P_y(z,t), \\ \frac{\partial H_z(z,t)}{\partial t} &= -\dot{\gamma} P_z(z,t).\end{aligned}\quad (47)$$

Обезразмерим дифференциальные уравнения (47), учитывая формулы (43), (7).

Получим безразмерные скалярные уравнения

$$\begin{aligned}\frac{\partial u_1(\bar{z}, \bar{t})}{\partial \bar{t}} &= G \frac{\partial^2 u_1(\bar{z}, \bar{t})}{\partial \bar{z}^2} - p_1(\bar{z}, \bar{t}), \quad \frac{\partial u_2(\bar{z}, \bar{t})}{\partial \bar{t}} = G \frac{\partial^2 u_2(\bar{z}, \bar{t})}{\partial \bar{z}^2} - p_2(\bar{z}, \bar{t}), \\ \frac{\partial u_3(\bar{z}, \bar{t})}{\partial \bar{t}} &= -p_3(\bar{z}, \bar{t}), \quad 0 < \bar{z} < 1, \quad 0 < \bar{t} < \infty,\end{aligned}\quad (48)$$

где $G = \frac{\tau}{\dot{\sigma}\mu_0\Delta^2}$.

Двухсторонние граничные условия

Обезразмерим заданную импульсную функцию $B_0(t')$ первичного поля (19)

$$B_0\left(\frac{t}{\tau}\right) = E_0 b(\bar{t}) = Z_0 H_0 b(\bar{t}), \quad (49)$$

где E_0 — максимальное значение напряжённости электрического поля $\vec{E}_0$, $b(\bar{t})$ — импульсная функция (20).

Приведём граничные условия (37), (30), (31) на плоскости $\Gamma_1(\bar{z}=0)$ к безразмерному виду, подставляя величины (43), (49). Получим двухсторонние граничные условия

$$\begin{aligned}\left(\frac{\partial u_1(\bar{z}, \bar{t})}{\partial \bar{z}} - K_0 u_1(\bar{z}, \bar{t})\right)\Big|_{\bar{z}=0} &= -2\dot{\sigma}\Delta Z_0 b(\bar{t}), \quad \left(\frac{\partial u_2(\bar{z}, \bar{t})}{\partial \bar{z}} - K_0 u_2(\bar{z}, \bar{t})\right)\Big|_{\bar{z}=0} = 0, \\ v_1\Big|_{\bar{z}=0} &= 0, \quad v_2\Big|_{\bar{z}=0} = 0, \quad v_3\Big|_{\bar{z}=0} = 0, \quad u_3\Big|_{\bar{z}=0} = 0,\end{aligned}\quad (50)$$

где $K_0 = \dot{\sigma}Z_0\Delta$.

Аналогично, преобразовывая граничные условия (42), (30), (31) на плоскости $\Gamma_2(\bar{z}=1)$, придём к безразмерным граничным условиям

$$\begin{aligned}\left(\frac{\partial u_1(\bar{z}, \bar{t})}{\partial \bar{z}} + K_0 u_1(\bar{z}, \bar{t})\right)\Big|_{\bar{z}=1} &= 0, \quad \left(\frac{\partial u_2(\bar{z}, \bar{t})}{\partial \bar{z}} + K_0 u_2(\bar{z}, \bar{t})\right)\Big|_{\bar{z}=1} = 0, \\ v_1\Big|_{\bar{z}=1} &= 0, \quad v_2\Big|_{\bar{z}=1} = 0, \quad v_3\Big|_{\bar{z}=1} = 0, \quad u_3\Big|_{\bar{z}=1} = 0.\end{aligned}\quad (51)$$

Исследуем импульсные электромагнитные поля, отражённые от экрана и прошедшие через экран D при воздействии первичного поля.

Эффективность экранирования

Воспользуемся двухсторонними граничными соотношениями (33) на плоскости Γ_1 . Учитывая (43) и (49), приведём соотношения (33) к физически безразмерному виду

$$H_0 u_1(0, \bar{t}) = \frac{1}{Z_0} (Z_0 H_0 b(\bar{t}) + B_1(\bar{t})), \quad H_0 u_2(0, \bar{t}) = \frac{1}{Z_0} A_1(\bar{t}).$$

В результате вычислим импульсную функцию отражённого от экрана электромагнитного поля (21)

$$A_1(\bar{t}) = E_0 u_2(0, \bar{t}), \quad B_1(\bar{t}) = E_0 (u_1(0, \bar{t}) - b(\bar{t})). \quad (52)$$

Двухсторонние граничные соотношения (39) на плоскости Γ_2 обезразмерим, учитывая (43).

Получим

$$H_0 u_1(1, \bar{t}) = \frac{1}{Z_0} B_2(\bar{t} - \bar{t}_0), \quad H_0 u_2(1, \bar{t}) = \frac{1}{Z_0} A_2(\bar{t} - \bar{t}_0), \quad \bar{t}_0 = \frac{\Delta}{c\tau}.$$

В результате вычислим импульсную функцию прошедшего через экран поля (23)

$$A_2(\bar{t}) = E_0 u_2(1, \bar{t} + \bar{t}_0), \quad B_2(\bar{t}) = E_0 u_1(1, \bar{t} + \bar{t}_0). \quad (53)$$

Коэффициент эффективности экранирования определим соотношением

$$\mathcal{E} = \frac{\max_{0 \leq \bar{t} < \infty} |\vec{E}_0(0, t)|}{\max_{0 \leq \bar{t} < \infty} |\vec{E}_2(\Delta, t)|}.$$

Вычислим модули электрических полей (19), (23)

$$\begin{aligned} |\vec{E}_0(0, t)| &= |B_0(\bar{t})| = E_0 b(\bar{t}), \quad \max_{0 \leq \bar{t} < \infty} |\vec{E}_0(0, t)| = E_0, \\ |\vec{E}_2(\Delta, t)| &= \sqrt{A_2^2(\bar{t} + \bar{t}_0) + B_2^2(\bar{t} + \bar{t}_0)} = E_0 \sqrt{u_2^2(1, \bar{t}) + u_1^2(1, \bar{t})}. \end{aligned}$$

Получим расчётную формулу для коэффициента эффективности экранирования

$$\mathcal{E} = \frac{1}{\max_{0 \leq \bar{t} < \infty} \sqrt{u_2^2(1, \bar{t}) + u_1^2(1, \bar{t})}}. \quad (54)$$

7. Начально-краевая задача экранирования для системы нелинейных уравнений для экрана из пермаллоя

Для системы скалярных дифференциальных уравнений (45), (48) сформулируем начально-краевую задачу с односторонними граничными условиями (50), (51), добавив нулевые начальные условия при $\bar{t} = 0$.

При решении начально-краевой задачи подлежат определению функции $u_1(\bar{z}, \bar{t})$, $u_2(\bar{z}, \bar{t})$, $u_3(\bar{z}, \bar{t})$, определяющие компоненты магнитного поля внутри слоя, и функции $v_1(\bar{z}, \bar{t})$, $v_2(\bar{z}, \bar{t})$, $v_3(\bar{z}, \bar{t})$, определяющие компоненты поля намагниченности пермаллоя в области $0 \leq \bar{z} \leq 1$, $0 \leq \bar{t} \leq 3$.

Математическая модель задачи. Требуется численно решить уравнения

$$\begin{aligned} \frac{\partial u_1(\bar{z}, \bar{t})}{\partial \bar{t}} &= G \frac{\partial^2 u_1(\bar{z}, \bar{t})}{\partial \bar{z}^2} - p_1(\bar{z}, \bar{t}), \quad \frac{\partial u_2(\bar{z}, \bar{t})}{\partial \bar{t}} = G \frac{\partial^2 u_2(\bar{z}, \bar{t})}{\partial \bar{z}^2} - p_2(\bar{z}, \bar{t}), \\ \frac{\partial u_3(\bar{z}, \bar{t})}{\partial \bar{t}} &= -p_3(\bar{z}, \bar{t}), \\ \frac{\partial v_1(\bar{z}, \bar{t})}{\partial \bar{t}} &= p_1(\bar{z}, \bar{t}), \quad \frac{\partial v_2(\bar{z}, \bar{t})}{\partial \bar{t}} = p_2(\bar{z}, \bar{t}), \quad \frac{\partial v_3(\bar{z}, \bar{t})}{\partial \bar{t}} = p_3(\bar{z}, \bar{t}), \end{aligned} \quad (55)$$

где $p_1(\bar{z}, \bar{t})$, $p_2(\bar{z}, \bar{t})$, $p_3(\bar{z}, \bar{t})$ определены в (46) при граничных условиях

$$\begin{aligned} \left(\frac{\partial u_1(\bar{z}, \bar{t})}{\partial \bar{z}} - K_0 u_1(\bar{z}, \bar{t}) \right) \Big|_{\bar{z}=0} &= -2K_0 b(\bar{t}), \quad \left(\frac{\partial u_2(\bar{z}, \bar{t})}{\partial \bar{z}} - K_0 u_2(\bar{z}, \bar{t}) \right) \Big|_{\bar{z}=0} = 0, \quad u_3(\bar{z}, \bar{t}) \Big|_{\bar{z}=0} = 0; \\ v_1(\bar{z}, \bar{t}) \Big|_{\bar{z}=0} &= 0, \quad v_2(\bar{z}, \bar{t}) \Big|_{\bar{z}=0} = 0, \quad v_3(\bar{z}, \bar{t}) \Big|_{\bar{z}=0} = 0, \quad 0 \leq \bar{t} \leq 3; \\ \left(\frac{\partial u_1(\bar{z}, \bar{t})}{\partial \bar{z}} + K_0 u_1(\bar{z}, \bar{t}) \right) \Big|_{\bar{z}=1} &= 0, \quad \left(\frac{\partial u_2(\bar{z}, \bar{t})}{\partial \bar{z}} + K_0 u_2(\bar{z}, \bar{t}) \right) \Big|_{\bar{z}=1} = 0, \quad u_3(\bar{z}, \bar{t}) \Big|_{\bar{z}=1} = 0; \\ v_1(\bar{z}, \bar{t}) \Big|_{\bar{z}=1} &= 0, \quad v_2(\bar{z}, \bar{t}) \Big|_{\bar{z}=1} = 0, \quad v_3(\bar{z}, \bar{t}) \Big|_{\bar{z}=1} = 0, \quad 0 \leq \bar{t} \leq 3, \end{aligned} \quad (56)$$

где $b(\bar{t})$ — безразмерная импульсная функция определенная в (20), с начальными условиями

$$\begin{aligned} u_1(\bar{z}, \bar{t}) \Big|_{\bar{t}=0} &= 0, \quad u_2(\bar{z}, \bar{t}) \Big|_{\bar{t}=0} = 0, \quad u_3(\bar{z}, \bar{t}) \Big|_{\bar{t}=0} = 0, \quad 0 \leq \bar{z} \leq 1; \\ v_1(\bar{z}, \bar{t}) \Big|_{\bar{t}=0} &= 0, \quad v_2(\bar{z}, \bar{t}) \Big|_{\bar{t}=0} = 0, \quad v_3(\bar{z}, \bar{t}) \Big|_{\bar{t}=0} = 0, \quad 0 \leq \bar{z} \leq 1. \end{aligned} \quad (57)$$

Для примера, в [5] выбраны материальные параметры пермаллоя $\dot{a} = 1.75 \cdot 10^{-6} \text{ м}^2$, $\dot{g} = 3.27 \cdot 10^{-6} \text{ м/А}$. Для остальных параметров задачи положим $H_0 = 10^3 \text{ А/м}$, $\tau = 10^{-3} \text{ с}$, $\Delta = 10^{-4} \text{ м}$, $\dot{\gamma} = 5.51 \cdot 10^2 \text{ м/(А·с)}$, $\dot{\sigma} = 1.71 \cdot 10^6 \text{ См/м}$, $Z_0 = 377 \text{ Ом}$, $\mu_0 = 1.26 \cdot 10^{-6} \text{ Гн/м}$. В этом случае безразмерные постоянные, входящие в постановку задачи, определяются величинами

$$G = \frac{\tau}{\dot{\sigma} \mu_0 \Delta^2} = 4.65 \cdot 10^4,$$

$$K_0 = \dot{\sigma} Z_0 \Delta = 6.45 \cdot 10^4,$$

$$F_1 = \frac{\dot{\gamma} \tau \dot{a} H_0}{\Delta^2} = 9.64 \cdot 10^4,$$

$$F_2 = \dot{\gamma} \tau H_0 = 5.51 \cdot 10^2,$$

$$F_3 = \dot{\gamma} \tau \dot{g} H_0^2 = 1.8.$$

После численной реализации решения начально-краевой задачи (55)-(57) вычисляются импульсные функции (52), (53), определяющие отражённое и прошедшее через экран поля (21), (23). Для численной оценки экранирующих свойств экрана из пермаллоя используется коэффициент эффективности экранирования (54) при различных значениях исходных параметров.

Список литературы

1. Виноградов А. П. Электродинамика композитных материалов. М.: Эдиториал УРСС, 2001. 206 с. ISBN: 978-5-8360-0283-1
2. Иванов О. В. Распространение электромагнитных волн в анизотропных и бианизотропных слоистых структурах. Ульяновск: УлГТУ, 2010. 262 с. ISBN 978-5-9795-0684-5
3. Неганов В. А., Осипов О. В. Современное состояние электродинамики искусственных киральных сред (обзор) // Физика волновых процессов и радиотехнические системы. 2005. Т. 8. № 1. С. 7-33.
4. Парфимович И. Д., Комаров Ф. Ф., Мильчанин О. С., Ткачев А. Г., Щегольков А. В. Микроволновые характеристики композитного материала на основе эпоксидного полимера с добавлением графеновых и ферритных наноматериалов // Доклады Национальной Академии Наук Беларуси. 2018. Т. 63. № 1. С. 22-28. DOI: 10.29235/1561-8323-2019-63-1-22-28
5. Ринкевич А. Б., Перов Д. В., Васильковский В. О., Лепаловский В. Н. Закономерности проникновения электромагнитных волн через металлические магнитные пленки // Журнал технической физики. 2009. Т. 79. № 9. С. 96-106.
6. Ерофеев В. Т. Математическая модель экранирования монохроматических электромагнитных полей плоскими экранами из пермаллоя // Информатика. 2019. Т. 16. № 2. С. 40-51.
7. Ерофеев В. Т., Козловская И. С. Аналитическое моделирование в электродинамике. М: Книжный дом «ЛИБРОКОМ», 2014. 304 с.
8. Кравченко В. Ф. Лекции по теории атомарных функций и некоторым их приложениям. Монография. М.: Радиотехника, 2003. 512 с. ISBN: 5-93108-019-8
9. Ерофеев В. Т., Громыко Г. Ф., Заяц Г. М. Экранирование побочных магнитных излучений кабеля пленочным цилиндрическим экраном с учетом нелинейных зависимостей // Журнал «Проблемы инфокоммуникаций», Минск: Белорусская государственная академия связи. 2019. № 1 (9). С. 62-70.
10. Ерофеев В. Т. Краевая задача дифракции пучков электромагнитных волн на плоском экране из биизотропных материалов // Известия НАН Беларуси. Сер. физ.-мат. наук. 2012. № 4. С. 72-79.
11. Бондаренко В. Ф., Ерофеев В. Т., Урбанович А. И. Моделирование экранирования электромагнитного излучения экранами с пространственной дисперсией, содержащими сферические частицы // Физические основы приборостроения. 2019. Т. 8. № 2(32). С. 91-101.
12. Ерофеев В. Т., Урбанович А. И. Моделирование прохождения гауссовых пучков электромагнитных волн через однослойный плоский экран из пермаллоя // Актуальные проблемы прочности: материалы международной научной конференции, Витебск, 25-29 мая 2020 года, под редакцией В. В. Рубаника. Молодежно: Типография «Победа», 2020. С. 151-153.
13. Ерофеев В. Т., Бондаренко В. Ф. Преобразование пучков электромагнитных волн при прохождении через экран из кирального метаматериала // Информатика. 2013. № 1. С. 5-17.
14. Ерофеев В. Т., Бондаренко В. Ф. Моделирование проникновения колеблющихся импульсных электромагнитных полей через экран, выполненный из сверхпроводящих материалов с временной дисперсией // Информатика. 2019. Т. 16. № 1. С. 46-58.

Поступила 19 июня 2020 г.

INITIAL BOUNDARY-VALUE PROBLEM OF SCREENING PULSED ELECTROMAGNETIC FIELDS BY PERMALLOY SHIELDS FOR NUMERICALL MODELING

V.T. EROFEENKO, A.I. URBANOVICH

doi: 10.25210/jfop-2002-016027

This paper presents the method to transform the three-domain nonlinear boundary-value problem of screening pulsed electromagnetic fields by a flat permalloy shield to the single-domain nonlinear boundary-value problem with the corresponding initial and one-sided boundary conditions. To simplify numerical realizations, the initial boundary-value problem is made physically dimensionless by the introduction of physically dimensionless time and dimensionless space coordinate.

In the first and the second sections, consideration is given to a system of vector nonlinear differential Maxwell equations which describe the propagation of electromagnetic fields and the fields of magnetization in permalloy. The boundary-value problem of screening is formulated in three domains: half-space of vacuum in front of the shield, flat permalloy layer, and half-space of vacuum behind the shield.

The third section presents a mathematical model for pulsed fields propagating in vacuum, with the energy flux orthogonal to the shield.

In the fourth and the fifth sections, the two-sided boundary conditions for interfacing at the shield planes are studied. The conditions of interfacing are transformed to the equivalent one-sided boundary conditions at the shield faces used to formulate a single-domain boundary-value problem.

The sixth section introduces the physically dimensionless components of a magnetic field and of a magnetization field. A system of vector equations is given in the form of a set of scalar physically-dimensionless parabolic equations. Single-sided boundary conditions at the shield are transformed to the first- and third-kind physically dimensionless conditions.

The seventh section presents the final mathematical model of the initial boundary-value problem for a permalloy shield in the form of a physically dimensionless system including six scalar parabolic equations with the boundary and initial conditions.

Уважаемые читатели!

Подписка на журнал оформляется через подписное агентство «Книга-Сервис» и объединенный каталог «Пресса России» (подписной индекс **29196**)

Журнал включен:

- В *Российский индекс научного цитирования (РИНЦ)*.
- В базу данных *Russian Science Citation Index (RSCI)* на платформе *Web of Science*.
- В *Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук*.

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ФИЗИЧЕСКИХ ПРОЦЕССОВ

УДК 004.922

ГЕНЕРАТОР ТЕСТОВЫХ ИЗОБРАЖЕНИЙ ДЛЯ ДЕТЕКТОРОВ УГЛОВ

© Авторы, 2020

doi: 10.25210/jfor-2002-028036

Егоров Д. П. — м.н.с., Институт радиотехники и электроники им. В.А.Котельникова Российской академии наук, Москва. SPIN: 8670–1629

Кравченко О. В. — н.с., Вычислительный центр им. А.А. Дородницына Федерального исследовательского центра «Информатика и управление» РАН, Москва. SPIN: 2349–9131

Митрофанова А. Ю. — студент кафедры «Высшая математика», Московский государственный технический университет им. Н. Э. Баумана, Москва. SPIN-код: 8931–2251

Чуриков Д. В. — с.н.с., Научно-технологический центр уникального приборостроения РАН, Москва; Институт радиотехники и электроники им. В.А.Котельникова РАН. SPIN: 2236–5778

Аннотация

В работе описан процесс разработки генератора простых геометрических фигур и синтез изображений вен для исследования модификаций детектора Харриса и более объективной оценки их качества, а также для уменьшения ресурсов, затрачиваемых на сбор рисунков кровеносной системы. Особое внимание уделено основным идеям и приёмам, используемым для создания генераторов: работе с цветом, определению координат возможных точек пересечения многоугольников, видимости углов, и аппроксимации кривых. Ключевые моменты пояснены иллюстрациями. Приведены примеры сгенерированных изображений.

Ключевые слова: детектор углов, генератор, цветовая модель, пересечение отрезков, видимость точки, аппроксимация кривых, биометрическая идентификация, синтетические данные, синтезатор тестовых данных

Abstract

The paper describes the development process of synthetic image generators (simple geometric shapes and veins) to study the modifications of the Harris detector and a more objective assessment of their quality, as well as to reduce the resources spent on collecting vein patterns. Particular attention is paid to the basic ideas and techniques used to create generators: working with color, determining the coordinates of possible intersection points of polygons, the visibility of corners, and approximating curves. Key points are illustrated. Examples of generated images are provided.

Keywords: corner detector, generator, color model, intersection of segments, point visibility, curve fitting, biometric identification, synthetic data, test data synthesizer

Введение

В настоящее время в связи с активным развитием цифровых технологий [1–21] приобретает большую актуальность задача защиты информации и разработка систем безопасности. Наибольшую надёжность, точность и удобство обеспечивают биометрические системы, так как основаны на уникальных индивидуальных для каждого человека параметрах. Среди таких систем наиболее распространёнными являются основанные на распознавании по отпечатку пальца, форме лица, сетчатке или радужной оболочке глаза. Однако эти методы обладают рядом недостатков. Во-первых, они основаны на внешних биометрических параметрах, что затрудняет процесс идентификации при загрязнении кожных покровов, порезах, влиянии мимики. Во-вторых, развитие технических средств даёт возможность компрометации и подделки внешних признаков (восстановление шаблона рисунка кожного покрова пальца по фотографии высокого разрешения, слепок отпечатка, созданию объёмных объектов и др.). В-третьих, более надёжные подходы доставляют неудобства: данные долго обрабатываются, во время сканирования (сетчатки) свет слепит глаза, а стоимость оборудования высока. Идентификация по рисунку венозной карты, предложенная ещё в 1992 году, позволяет исключить эти недостатки. Рисунок вен [22] формируется благодаря поглощению гемоглобином в крови инфракрасного излучения, а сама технология является бесконтактной, что с одной стороны снижает психологический барьер и является гигиеничным, а с другой позволяет предпринимать простые меры по защите биометрической информации.

В настоящее время популярность этого метода возрастает. Встречается идентификация по рисункам вен пальцев, тыльной стороны ладони, запястья [1,2]. Для самой идентификации используют нейронные сети [3] или детекторы углов [4,5] — методы извлечения из изображений одной из разновидностей особых точек, т.е. такой точки m изображения, окрестность которой $U_o(m)$ можно отличить от окрестности любой другой точки изображения $U_o(n)$ в некоторой другой окрестности особой точки $U_o(m)$. Следует отметить, что существует большое количество видов детекторов углов, таких как детектор

Харриса [6], алгоритм Smallest Univalve Segment Assimilation Nucleus (SUSAN) [7], детектор Фёрстнера [8] и др. Также существует и множество их модификаций, устраняющих некоторые из недостатков базовых алгоритмов (например, [9,10]).

На пути к реализации решения задачи идентификации по рисункам венозной системы стоят следующие несколько проблем: выбор подходящего детектора (или его модификации), его тестирование, обучение нейросети. Для этого необходимы базы данных изображений, а чтобы проверить качество работы модифицированного детектора, обычно используют библиотеки реальных изображений (например, [11]). Однако в этом методе есть существенный недостаток — оператор базы данных не может точно определить координаты углов на реальном изображении, поэтому в оценку результатов работы детектора закладывается априорная погрешность. Можно найти также и уже готовые базы данных с рисунками вен [12], но доступ к ним, как правило, ограничен, а полученные изображения необходимо дополнительно обработать для дальнейшего тестирования. Поэтому не всегда можно воспользоваться существующими базами данных, а для того, чтобы собрать их самостоятельно, потребуется время, волонтеры и оборудование.

Таким образом, приходим к проблеме генерации тестовых изображений. В данной работе во избежание описанных выше трудностей предлагается следующая схема. Для тестирования детекторов углов будем генерировать изображения простых геометрических фигур — многоугольников. Такие тестовые данные будут иметь заранее известные координаты углов и, соответственно, ориентацию. Это не исключает необходимости использования реальных тестовых изображений, но позволит объективно оценить работу детектора. Чтобы проверить работу детектора на рисунках карт вен, а в дальнейшем обучить нейронную сеть, необходимо перейти от генерации простых фигур к более сложным и приближенным к реальности изображениям.

Разработка приложений

Оба приложения разработаны на языке Python версии 3.7. Для создания генераторов использовались библиотеки Tkinter для создания интерфейса, Pillow для работы с изображениями, библиотеки NumPy и Random для математических вычислений и генерации случайных величин, а также библиотека Wand для улучшения качества создаваемых изображений. Помимо этого, для синтеза рисунков вен пальца использовалась библиотека SciPy для аппроксимации кривых. Подробнее об этом будет рассказано ниже.

Рис. 1. Алгоритм генерации изображения.

В разработанном генераторе простых геометрических фигур можно создать до двадцати изображений с заданными параметрами размера и количества фигур на изображении. Алгоритм генерации одного изображения приведён на рис. 1, а основные используемые переменные приведены в табл. 1.

Таблица 1. Основные используемые переменные

Переменная	Идентификатор	Тип данных	Область допустимых значений
Ширина изображения	ImSizeW	целое	[64;1024]
Высота изображения	ImSizeH	целое	[64;1024]
Радиус описанной окружности	r	целое	$\left[\sqrt{0,1 \frac{S}{\pi}}; \sqrt{0,3 \frac{S}{\pi}} \right]$, где S — площадь изображения
Угол поворота	angle	целое	[0° ;360°]
Количество вершин	n	целое	[3;6]
Х-координата центра многоугольника	center_x	целое	[5;ImSizeW – 5] или $\left[\frac{\text{ImSizeW}}{2} - \frac{\text{ImSizeW}}{18}; \frac{\text{ImSizeW}}{2} + \frac{\text{ImSizeW}}{18} \right]$
Y-координата центра многоугольника	center_y	целое	[5;ImSizeH – 5] или $\left[\frac{\text{ImSizeH}}{2} - \frac{\text{ImSizeH}}{18}; \frac{\text{ImSizeH}}{2} + \frac{\text{ImSizeH}}{18} \right]$
Расстояние между двумя описанными окружностями	d	вещественное	$\left[0; \sqrt{2(100 - \text{ImSizeH} - \text{ImSizeW}) + \text{ImSizeH}^2 + \text{ImSizeW}^2} \right]$
Координаты всех видимых углов	corners	список кортежей	—

Подробнее остановимся на ключевых моментах приложения.

Работа с цветом

Рассмотрим работу приложения с цветом. Наибольший интерес представляет тестирование детекторов на изображениях разного цвета. Предлагается незначительно изменять выбранный пользователем цвет на каждом изображении. Проще всего это было сделать в пространстве HSV [13], где было необходимо изменять лишь элементы s и v , отвечающие за чистоту (насыщенность) цвета и за его яркость, а элемент h оставлять неизменным из-за того, что он отвечает непосредственно за значение цвета, т.е. цветовой тон.

Учитывая описанные выше характеристики, а также некоторые особенности работы с изображениями в Python, алгоритм преобразования цвета выглядит следующим образом:

1. получить RGB компоненты выбранного пользователем цвета;
2. перевести выбранный цвет в пространство HSV;
3. выбрать случайные значения s и v из заданного интервала;
4. перевести случайно полученный цвет в RGB-пространство;
5. преобразовать полученный цвет в строку с HEX значением цвета.

Для перевода цвета в HSV-пространство следует разделить значение каждого из RGB компонентов на 255, а для обратного перехода в RGB из HSV умножить значение компоненты на то же число. Для эле-

ментов s и v в зависимости от их значений были определены следующие возможные интервалы, в пределах которых они могли изменяться:

$$s = \begin{cases} [s; s+0,1], s-0,1 \leq 0, \\ [s-0,1; s], s+0,1 \geq 1, \\ [s-0,1; s+0,1], 0 < s-0,1 < s+0,1 < 1, \end{cases}$$

$$v = \begin{cases} [v; v+0,2], v-0,2 \leq 0, \\ [v-0,2; v], v+0,2 \geq 1, \\ [v-0,2; v+0,2], 0 < v-0,2 < v+0,2 < 1. \end{cases}$$

Далее, рассмотрим основные этапы синтеза изображения, содержащего геометрические фигуры.

Определение центра многоугольника

При определении центра многоугольника необходимо учитывать тот факт, что многоугольник может не поместиться на изображении целиком. Здесь следует учитывать, хочет ли пользователь видеть в своей тестовой выборке подобные изображения. На рис. 2 приведены примеры расположения описанных вокруг многоугольников окружностей. На этом рисунке внешний квадрат является самым холстом, а внутренний квадрат представляет собой рамку, в пределах которой располагается центр окружности.

Рис. 2. Примеры положения центров описанных окружностей.

Определение координат углов многоугольника

После определения центра многоугольника (с координатами x_0, y_0), вычисляются координаты его углов. Сначала фиксируются координаты первого угла многоугольника, который всегда находился в верхней точке описанной окружности. Затем по формуле определяются координаты искомой точки x', y'

$$\begin{aligned} x' &= x_0 + (x - x_0)\cos\alpha - (y - y_0)\sin\alpha, \\ y' &= y_0 + (x - x_0)\sin\alpha + (y - y_0)\cos\alpha, \end{aligned} \quad (1)$$

где x, y — координаты предыдущей найденной точки, α — угол поворота, для правильных многоугольников определяемый формулой

$$\alpha = \frac{360}{n}. \quad (2)$$

После этого уже полученная точка поворачивалась вокруг центра на тот же угол α .

Для определения координат углов неправильных многоугольников вместо поворота на один и тот же угол α точки поворачивались на произвольные углы из интервалов, сумма точных верхних границ которых не превышала 345° (эмпирическое значение).

После этого как для правильных, так и для неправильных многоугольников выполнялось преобразование полученных координат по следующей формуле:

$$\begin{aligned} x'' &= x_0 + (x - x_0)\cos\varphi - (y - y_0)\sin\varphi, \\ y'' &= y_0 + (x - x_0)\sin\varphi + (y - y_0)\cos\varphi, \end{aligned} \quad (3)$$

где φ — случайно сгенерированный угол поворота в радианах, x, y — старые координаты углов, x'', y'' — новые координаты углов.

Определение координат пересечения сторон выпуклых многоугольников

В связи с тем, что в приложении предусмотрена возможность построения нескольких геометрических фигур на одном изображении, был решен вопрос об определении координат их пересечений. Для проверки пересечения сторон многоугольников, пример приведён на рис. 3, применялся простой в реализации метод, предложенный в [14].

Рис. 3. Пример пересекающихся отрезков.

Возвращаясь к задаче генерирования синтетического изображения геометрических фигур, имеем два вложенных цикла, в которых попарно проверяются каждые из отрезков, образующих многоугольник. В случае их пересечения находим координаты точки, в которой они пересекаются. Решая систему, сохраняем полученные координаты в список углов *corners*. Использование векторного произведения для определения пересечения отрезков является простым способом, однако для его применения в задаче определения точек пересечения многоугольников требуется найти от двух до четырёх векторных произведений для всех возможных пар сторон геометрических фигур. Существуют различные алгоритмы для решения поставленной задачи [15]. Среди них можно выделить, например, алгоритм О'Рурка, суть которого заключается в том, чтобы не двигаться по той границе многоугольника, текущее ребро которой еще может содержать необнаруженную точку пересечения. В силу того, что в рассматриваемой задаче многоугольники являются выпуклыми, то для неё возможно применить алгоритм О'Рурка.

Заметим, что синтезирование невыпуклых многоугольников возможно, и подобная задача может быть решена в рамках подхода, изложенного в [18–21]. В этом случае, необходимо перейти от координатного представления вершин отрезков к так называемому комплексному представлению отрезков, соединяющих вершины. Далее, составляя характеристическую функцию области (R-функцию), принимающую положительные значения внутри области, нулевые на границе и отрицательные вне области, можно однозначно сделать вывод о положении той или иной вершины относительно многоугольника. При этом, рассуждения, изложенные выше остаются справедливыми и в задаче синтеза, когда искомым многоугольником формируется добавлением новых вершин. Однако данный алгоритм без дополнительной оптимизации является трудоёмким по времени, и уступает в скорости выполнения более простым геометрическим методам, хотя и превосходит последние в универсальности.

Определение видимости угла многоугольника

Если расстояние между центрами описанных возле многоугольников окружностей меньше суммы их радиусов, то, помимо пересечения многоугольников, возможно, что одного из углов (нескольких или всех) геометрической фигуры не видно за второй. В этом случае нужно определить координаты таких углов и в список *corners* их не заносить.

Чтобы понять, видно угол первого многоугольника за вторым или нет, нужно определить, находится ли точка с координатами этого угла внутри или снаружи многоугольника. Для этого был использован метод трассировки лучей [16]. Этот метод заключается в том, что из проверяемой точки выпускается луч в произвольном или предопределённом направлении. Если точка лежит внутри многоугольника, то луч пересечёт многоугольник нечётное количество раз. Если же точка находится снаружи, то количество пересечений будет чётным.

В случае касания углов, как было замечено ранее, считается, что стороны многоугольников пересекаются. Поэтому список углов `corners` пополняется координатой этого пересечения на предыдущем этапе. Перед внесением координат углов геометрических фигур в список `corners` также осуществляется проверка на их наличие в этом списке. Поэтому вне зависимости от того, определит трассировка лучей угол как видимый, или как невидимый (в зависимости от расположения фигур может встретиться как один, так и другой вариант), угол всё равно войдёт в итоговый список, причём только один раз.

Синтезирование рисунков вен

После разработки генератора простых геометрических фигур был создан генератор более сложных, приближенных к реальным, изображений. Было решено сузить задачу до генерации рисунков вен указательного пальца правой руки. Для генерации таких изображений потребовалось выполнить следующие шаги:

1. найти достоверное изображение кровеносной системы;
2. разбить вены указательного пальца на изображении на участки;
3. создать приложение, с помощью которого можно обвести каждый из участков и сохранить координаты нарисованной линии в таблицу;
4. для каждого участка подобрать функцию, которая наилучшим образом аппроксимирует полученные координаты;
5. для каждого участка найти интервал, в котором могут изменяться найденные параметры функций;
6. добавить изменение толщины сосудов, угла поворота изображения и размытия.

Таким образом, сначала нужно создать некоторый шаблон кровеносной системы, и только потом приложение, способное генерировать изображения с изменениями в шаблоне.

Создание шаблона рисунка вен

В качестве опорного изображения был выбран рисунок поверхностных вен предплечья и кисти правой верхней конечности, взятое из третьего тома медицинского атласа Синельникова [22]. На этом изображении выбран и увеличен фрагмент указательного пальца, после чего вены на нём разбиты на 24 участка. После этого проведен процесс обводки этих участков и занесения координат в таблицу.

Для аппроксимации полученных результатов выбраны многочлены не выше шестой степени. Оптимальные параметры a, b, c, d, e, f, g функций

$$\begin{aligned} & ax + b, \\ & ax^2 + bx + c, \\ & ax^4 + bx^3 + cx^2 + dx + e, \\ & ax^6 + bx^5 + cx^4 + dx^3 + ex^2 + fx + g \end{aligned} \quad (4)$$

определялись с помощью функции `curve_fit` из библиотеки `Scipy`. Результат такого приближения можно видеть на рис. 4, на котором точками указаны координаты положения курсора при обводке кровеносного сосуда.

Рис. 4. Пример аппроксимации одного из участков вен полиномами разных степеней.

Заключение

В работе предложены алгоритмы генераторов простых фигур, а также изображений карт вен на основе простых цветовых моделей, аффинных преобразований [17], и трассировки лучей. Основной моделью для синтеза изображений карт вен является модель аппроксимации кривых алгебраическими полиномами не выше 6 порядка. На основе описанных алгоритмов были разработаны компьютерные приложения, позволяющие пользователю проводить синтез тестовых изображений и изображений карт вен. Полученные синтетические изображения представлены на рис. 5–7.

Рис. 5. Примеры синтетических изображений геометрических фигур.

Рис. 6. Сгенерированные рисунки вен.

Рис. 7. Различия сгенерированных рисунков вен.

Заметим, что рисунки венозной системы имеют ряд недостатков, таких как плохое качество отрисовки и низкая точность. Для повышения точности генерируемых изображений можно прибегнуть к следующим методам:

1. вместо применения тачпада для обводки участков вен прибегнуть к использованию графического планшета и пера, что повысит точность нарисованного шаблона и, следовательно, аппроксимации;
2. увеличить количество обводок одного и того же участка вены и аппроксимировать средние значения;
3. скорректировать интервалы изменения параметров функций.

Несмотря на ряд недостатков в генерации рисунков кровеносной системы, эти изображения могут быть использованы для тестирования детектора Харриса (или других алгоритмов поиска особых точек) в поиске углов на рисунках венозных карт. В работе предлагается 2 различных генератора изображений, один из которых создаёт изображения простых геометрических фигур и может быть использован в исследовании качества детекторов углов и их модификаций, а другой синтезирует рисунки вен, с помощью чего появляется возможность тестирования детектора углов на близкой к реальной задаче и, тем самым, пополнить выборку обучающих данных для нейронной сети сгенерированными изображениями. Отметим, что описанный алгоритм синтеза изображений вен можно применить также и к генерации рисунков артерий и других кровеносных и лимфатических сосудов. Помимо использования сгенерированных изображений сосудов в области идентификации, они также могут быть применены в соответствующей сфере медицины — ангиологии. Реализованный алгоритм может быть полезен для обучения нейронных сетей, способных диагностировать заболевания сосудов, что является продолжением данной работы и служит предметом дальнейших исследований.

Работа выполнена при финансовой поддержке РФФИ (грант № 18-29-02108).

Список литературы

1. Kulkarni, S., Pandit, M. Biometric recognition system based on dorsal hand veins // Int J Innov Res Sci Eng Technol. 2016. Vol. 5. No. 9. P. 18899–18905.
2. Sun, S., Li, S., and Guo, Z. Binary Filter for Fast Vessel Pattern Extraction // Neural Processing Letters. 2019. Vol. 49. No. 3. P. 979–993.
3. Hong, H.G., Lee, M.B., and Park, K.R. Convolutional neural network-based finger-vein recognition using NIR image sensors // Sensors. 2017. Vol. 17. No. 6. P. 1297.
4. Детекторы углов, 2014, URL: <https://habr.com/ru/post/244541/> (дата обращения: 20.02.2020)
5. Cui, J. et al. Corners detection on finger vein images using the improved Harris algorithm // Optik. 2014. Vol. 125. No. 17. P. 4668–4671.
6. Harris, C. G. et al. A combined corner and edge detector // Alvey vision conference. 1988. Vol. 15. No. 50. P. 10–5244.
7. Tuytelaars, T. et al. Local invariant feature detectors: a survey // Foundations and trends® in computer graphics and vision. 2008. Vol. 3. No. 3. P. 177–280
8. Furstner W., Gblich E. A fast operator for detection and precise location of distinct points, corners and centres of circular features // Proc. ISPRS intercommission conference on fast processing of photogrammetric data. 1987. P. 281–305.
9. Pei, S. C., Ding, J. J. Improved harris' algorithm for corner and edge detections // IEEE International Conference on Image Processing. IEEE. 2007. Vol. 3. P. III-57-III-60.
10. Mahmoodi, S., Gunn, S. Scale space smoothing, image feature extraction and Bessel filters // Scandinavian Conference on Image Analysis. Springer, Berlin, Heidelberg, 2011. P. 625–634.
11. CVonline: Image Databases, URL: <http://homepages.inf.ed.ac.uk/rbf/CVonline/Imagedbase.htm> (дата обращения: 20.02.2020)
12. Idiap Dataset Distribution Portal URL: <https://www.idiap.ch/dataset/vera-fingervein>, (дата обращения: 20.02.2020)
13. Поиск объекта по цвету. Цветовое пространство HSV, URL: <http://robocraft.ru/blog/computervision/402.html> (дата обращения: 20.02.2020)
14. Простой алгоритм определения пересечения двух отрезков, 2015, URL: <https://habr.com/ru/post/267037/> (дата обращения: 20.02.2020)
15. Ченцов О.В., Скворцов А.В. Обзор алгоритмов построения оверлеев многоугольников // Вестн. Том. гос. ун-та. 2003. № 280. URL: <https://cyberleninka.ru/article/n/obzor-algoritmov-postroeniya-overleev-mnogougolnikov> (дата обращения: 20.02.2020).
16. Методы определения принадлежности точки многоугольнику, 2016, URL: <https://habr.com/ru/post/301102/> (дата обращения: 20.02.2020)
17. Кудрина М.А., Мурзин А.В. Аффинные преобразования объектов в компьютерной графике // НиКа. 2014. URL: <https://cyberleninka.ru/article/n/affinnye-preobrazovaniya-obektov-v-kompyuternoy-grafike> (дата обращения: 20.02.2020)
18. Кравченко В. Ф., Чуриков Д. В. Анализ временных рядов комплексными ВА-системами функций Кравченко // Электромагнитные волны и электронные системы. 2010. Т. 15. № 7. С. 3–17.
19. Кравченко В. Ф., Кравченко О. В., Чуриков Д. В. Применение R-функций, атомарных и ВА-систем функций в информационных технологиях. Обзор // Актуальные проблемы современного образования. 2018. Т. 2. С. 13–23.
20. Кравченко В. Ф., Чуриков Д. В., Юрин А. В. Аналитическое описание локусов сложной формы r-операциями и атомарными функциями. Цифровая обработка сигналов и изображений // Успехи современной радиоэлектроники. 2007. № 3. С. 6–37.

21. Кравченко В. Ф., Чуриков Д. В. Цифровая обработка сигналов атомарными функциями и вейвлетами. М.: Техносфера, 2019. 182 с. ISBN 978-5-94836-506-0
22. Синельников Р. Д., Синельников Я. Р., Синельников А. Я. Атлас анатомии человека. В 4-х томах. Том 3. Учение о сосудах и лимфоидных органах. М.: Новая волна, 2020. 216 с. ISBN: 978-5-7864-0201-9

Поступила 17 марта 2020 г.

IMAGE GENERATOR FOR TESTING CORNER DETECTORS

D. P. EGOROV, O. V. KRAVCHENKO, A. YU. MITROFANOVA, AND D. V. CHURIKOV

doi: 10.25210/jfop-2002-028036

The paper describes the development process of synthetic image generators which are simple geometric shapes and veins to study the modifications of the Harris detector and a more objective assessment of their quality, as well as to reduce the resources spent on collecting vein patterns. Particular attention is paid to the basic ideas and techniques used to create generators: working with color, determining the coordinates of possible intersection points of polygons, the visibility of corners, and approximating curves. Key points are illustrated. Examples of generated images are provided.

Уважаемые читатели!

Подписка на журнал оформляется через подписное агентство «Книга-Сервис» и объединенный каталог «Пресса России» (подписной индекс **29196**)

Журнал включен:

- В *Российский индекс научного цитирования (РИНЦ)*.
- В базу данных *Russian Science Citation Index (RSCI)* на платформе *Web of Science*.
- В *Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук*.

НОВЫЕ МАТЕМАТИЧЕСКИЕ И ФИЗИЧЕСКИЕ МЕТОДЫ

УДК 007.001.33

ЯЗЫК ЛОГИКИ ПРЕДИКАТОВ В СИСТЕМАХ
ОБРАБОТКИ ИНФОРМАЦИИ В БАЗАХ ЗНАНИЙ

© Авторы, 2020

doi: 10.25210/jfor-2002-037047

Болотова Е. Е. — магистр, кафедра «Высшая математика», Московский государственный технический университет им. Н.Э. Баумана. E-mail: lisayshka_bolotova@mail.ru**Бутенко Ю. И.** — к.т.н., доц. кафедры «Романо-германские языки», Московский государственный технический университет им. Н.Э. Баумана, Москва. E-mail: iuliiabutenko2015@yandex.ru**Сидняев Н. И.** — д.т.н., проф., кафедра «Высшая математика», Московский государственный технический университет им. Н.Э. Баумана. E-mail: Sidn_ni@mail.ru**Аннотация**

В статье раскрыты основные принципы построения интеллектуальных систем по инженерии знаний. Представлены исследования по искусственному интеллекту с использованием нейросистем и методов представления знаний в экспертных системах. Обсуждены логика предикатов, синтаксис и семантика языка предикатов, правила выводов в логике предикатов, а также способы представления информации в компьютерных системах на различных этапах решения проблемы, модели решения и классификации проблем. Рассмотрены основные структуры систем обработки знаний и языков представления знаний. Разработана структурно-функциональная схема интеллектуальной системы для нейрокомпьютерной реализации конструктивных оперативно-советующих экспертных систем. Показано, что ядром любой интеллектуальной системы является база знаний, которая описывает определенную предметную область. Постулируется, что выбор правильной модели знания в экспертных системах играет важное значение, и в большинстве случаев становится основой выбора подхода к решению той или иной задачи.

Ключевые слова: базы данных, логика предикатов, язык, экспертные системы, структуры, знания

Abstract

The article reveals the basic principles of building intelligent systems for knowledge engineering. Research on artificial intelligence using neurosystems and methods of knowledge representation in expert systems is presented. The logic of predicates, syntax and semantics of predicate language, rules of inference in predicate logic, as well as ways of presenting information in computer systems at various stages of problem solving, models of solving and classification of problems are discussed. The basic structures of knowledge processing systems and knowledge representation languages are considered. The structural and functional scheme of the intellectual system for the neurocomputer implementation of constructive operational-advising expert systems is developed. It is shown that the core of any intellectual system is a knowledge base that describes a specific subject area. It is postulated that the choice of the correct model of knowledge in expert systems plays an important role, and in most cases becomes the basis for choosing an approach to solving a particular problem.

Keywords: database, predicate logic, language, expert systems, structures, knowledge

Введение

С развитием систем искусственного интеллекта появилась потребность в представлении знаний о различных предметных областях с целью эффективной обработки этих знаний с помощью ЭВМ. Потребовались информационные системы, которые могли бы использовать большое количество знаний для решения сложных задач, а также появилась необходимость в создании баз знаний для описания различных предметных областей. Одним из направлений в исследованиях по искусственному интеллекту, изучающим способы описания объектов реального мира, является представление знаний. Существует множество различных моделей представлений знаний (см. рис. 1), однако в области обработки и представления знаний все большую популярность набирает логика предикатов как один из мощнейших инструментов вывода информации. Основная идея при построении логических моделей знаний заключается в следующем: вся информация, необходимая для решения прикладных задач, рассматривается как совокупность фактов и утверждений, которые представляются как формулы в некоторой логике. Знания отображаются совокупностью таких формул, а получение новых знаний сводится к реализации процедур логического вывода. В качестве «фундамента» в логических моделях используется классический аппарат математической логики, методы которой достаточно хорошо изучены и формально обоснованы. Язык математической логики позволяет формально и в то же время доступно описывать понятия предметной области и связи между ними. Однако независимо от того, используется или не используется логика предикатов для представления знаний, она составляет теоретический фундамент в систематике описания, а по своему существу находится в основе всех представлений знаний [1].

Рис. 1. Модели представления знаний.

Представление знаний будет эквивалентно логике предикатов и полностью унаследует все ее свойства, если только определить язык представления знаний через метод представления, полностью соответствующий логике предикатов. Например, формализм ЕСЛИ — ТО продукционных правил по виду аналогичен отношению импликации в логике предикатов; что же касается дуг, которые связывают узлы в семантической сети, то если узлу поставить в соответствие терм, а дуге — предикатный символ, непосредственно получаем предикат [2]. Например, для представления вида «ЕСЛИ А, ТО В» выполняется антитезис «ЕСЛИ не В, ТО не А» ($\text{ЕСЛИ } \sim B, \text{ ТО } \sim A$). Следовательно, исходя из знаний вида «В» и «ЕСЛИ А, ТО В», на вопрос « $\sim A$?» должен последовать ответ «Да». Когда проводится формальный вывод методами символьной обработки, то такой результат получить нельзя, если положить, что А и $\sim A$ — различные объекты, однако до какой степени это соответствие учитывается в современных системах инженерии знаний? Если ответ на вопрос дается методом, при котором в выводе предполагается направленность, как это имело место в восходящих или нисходящих выводах, то для одного правила должны быть заданы все его эквивалентные выражения. Например, когда имеется правило «ЕСЛИ $A_1, A_2, \dots, A_n$, ТО В», то в качестве эквивалентных ему выражений должны выполняться правила

$$\text{«ЕСЛИ } B, A_2, \dots, A_n \text{ ТО } \sim A_1\text{»}, \text{«ЕСЛИ } A_1, \sim B, \dots, A_n \text{ ТО } \sim A_2\text{»}, \dots, \text{«ЕСЛИ } A_1, A_2, \dots, \sim B \text{ ТО } \sim A\text{»}$$

или же эти эквивалентные выражения должны создаваться внутри механизма выводов. После создания исходная формула будет преобразована к виду: $\sim A_1 \vee \sim A_2 \vee \dots \vee \sim A_n \vee B$. Это не что иное, как использование резолюции при доказательстве. Все это вызвано тем, что полнота метода резолюции нивелирует в импликационном отношении внешнее различие, базирующееся на положении антецедента и консеквента внутри логического выражения, а сам метод может использовать только ту информацию, которая семантически содержится внутри этого выражения. В логике предикатов первого порядка основными элементами, с которыми манипулируют синтаксические правила, являются термы и предикаты. Терм представляет сущности описываемого явления, т.е. те объекты, которые уже не разделяются на более мелкие фрагменты, а предикаты содержат в качестве своих аргументов только сущности подобного рода и описывают атрибуты сущностей и их отношения (Рис. 3). Вдобавок использование переменных допускается только в термах. Однако в естественном языке содержатся выражения, которым нельзя найти соответствия, если не превратить в переменные сами предикаты. Например, используемый в качестве примера предикат МОДЕЛИ ПРЕДСТАВЛЕНИЯ ЗНАНИЙ (х, у) описывает «х и у, которые находятся в отношении МОДЕЛИ ПРЕДСТАВЛЕНИЯ ЗНАНИЙ». Однако, если формулой языка предикатов должно описываться выражение, которое на естественном языке соответствует вопросительному предложению «В каких отношениях находятся фреймовая модель и семантическая сеть?», то тогда сам предикатный символ должен быть сделан переменной ($\exists z$)z (Фреймовая модель, семантическая сеть)? Можно так расширить синтаксические правила, что они уже будут допускать подобного рода превращение предикатных символов в переменные и вновь созданная таким образом система описания превратится в расширенную систему, которая будет включать в себя все предложения, допускаемые синтаксическими правилами логики предикатов первого порядка. Такая система будет называться логикой предикатов второго порядка. Если определять более строго, то логикой предикатов второго порядка является система описа-

ния, которая может включать в качестве аргументов предикатов не только термы, описывающие сущности, но и предложения предикатов первого порядка. Рассмотрим в соответствии с приведенным выше примером предикат ОТНОШЕНИЕ (x, y, z) . Положим, этот предикат говорит о том, что «отношение между y и z есть x », и пусть МОДЕЛИ ПРЕДСТАВЛЕНИЯ ЗНАНИЙ (y, z) , являющийся предикатом первого порядка, входит в него в виде ОТНОШЕНИЕ (МОДЕЛИ ПРЕДСТАВЛЕНИЯ ЗНАНИЙ, y, z). Если в этом случае записать $(\exists x)$ ОТНОШЕНИЕ $(x, \text{фреймовая модель, семантическая сеть})$, то эта предикатная формула описывает предыдущее вопросительное предложение.

В приведенном выше примере ОТНОШЕНИЕ (x, y, z) существует множество самых различных отношений, которые могут возникать между y и z , куда относятся отношения «Модели представления знаний и аппарат семантических сетей», «Системы искусственного интеллекта и их подсистемы» и др. Допустим, что все это разнообразие заранее ограничено определенным числом отношений, помимо которых других отношений не существует. В этом случае предикат второго порядка можно разложить на группу предикатов первого порядка. Например, допустим, что при подстановке отношений в x получаем ОТНОШЕНИЕ (Модели представления знаний, y, z), ОТНОШЕНИЕ (Системы искусственного интеллекта, y, z), ... Затем это перепишем в виде ОТНОШЕНИЕ- КЛАССИФИКАЦИЯ МОДЕЛЕЙ ПРЕДСТАВЛЕНИЯ ЗНАНИЙ (y, z) , ОТНОШЕНИЕ-КЛАССИФИКАЦИЯ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА (y, z) , ... Это все есть предикаты первого порядка и к тому же их число конечно. Тем самым внутри выражения, относящегося к ОТНОШЕНИЕ (x, y, z) , можно получить разложение выражения по переменной x в виде

$$\begin{aligned} &(\forall x) \text{ОТНОШЕНИЕ}(x, y, z) = \text{ОТНОШЕНИЕ- КЛАССИФИКАЦИЯ МОДЕЛЕЙ ПРЕДСТАВЛЕНИЯ ЗНАНИЙ} \\ &\quad (y, z) \wedge \text{ОТНОШЕНИЕ- КЛАССИФИКАЦИЯ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА}(y, z) \wedge \dots \\ &(\exists x) \text{ОТНОШЕНИЕ}(x, y, z) = \text{ОТНОШЕНИЕ- КЛАССИФИКАЦИЯ МОДЕЛЕЙ ПРЕДСТАВЛЕНИЯ ЗНАНИЙ} \\ &\quad (y, z) \vee \text{ОТНОШЕНИЕ- КЛАССИФИКАЦИЯ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА}(y, z) \vee \dots \end{aligned}$$

Таким образом, становится возможным описание правил использования предикатов второго порядка. Если такое преобразование допустимо, то при подобном подходе описательные возможности предложений будут достаточно велики и в рамках логики предикатов первого порядка.

Структуры данных и язык предикатов

В языках декларативного типа семантика текста конструируется из семантики отдельных слов и грамматических правил. Семантика же отдельных слов является основным неделимым семантическим элементом в понимании. Таким же образом и семантика, описывающая общую структуру, может быть выражена основными элементами структуры и формулами структуризации [3–4]. Подобного рода базовые структурные элементы имеют собственный, присущий только им смысл, и их нельзя расчленить на более мелкие элементы.

Проблема состоит в таком формальном определении множества, при котором обеспечивается формализация логики предикатов в качестве системы описания. Само по себе множество — это совокупность объектов, однако интуитивным методом, который связывает области, имеющие некоторые присущие им свойства, в единое целое через выраженные предикатами аксиомы, является гипотеза о существовании множества, состоящего из всех объектов, обладающих этими свойствами. Обычно такие свойства могут быть какими угодно и будут далее обозначаться предикатом $P(x)$. В таком случае для произвольного свойства $P(x)$ существует множество y , для которого $x \in y \leftrightarrow P(x)$. Поскольку $P(x)$ — произвольное свойство, то можно положить, что $X \notin X$. Пусть R — множество, определенное таким свойством его элементов, а именно: $x \in R \leftrightarrow x \notin x$. Таким образом, R , которое определено здесь совершенно формально, будет представлять из себя, если это выразить словесно, «множество, состоящее из самого множества, которое не является собственным элементом». Однако приняв, что такое множество существует, попробуем задать вопрос: «является ли R элементом R ?». Положительного ответа на этот вопрос не существует, поскольку выполняется $R \in R \leftrightarrow R \notin R$, а это является очевидным противоречием.

Здесь необходимо рассмотреть другие подходы к определению множества, которые являются, насколько возможно, общими и в то же время не приводят к ситуациям, характерным для обратной теоремы.

Необходимо привести несколько фундаментальных аксиом:

$$(1) \text{Аксиома общности: } (\forall x)(\forall y)[x = y \leftrightarrow (\forall z)[z \in x \leftrightarrow z \in y]].$$

Два множества с достоверно одинаковыми элементами равны между собой.

$$(2) \text{Аксиома включения: } (\forall u)(\exists v)(\forall x)[x \in v \leftrightarrow x \in u \wedge P(x)].$$

Если заданы множество u и формула логики предикатов $P(x)$, то существует множество v , содержащее все те элементы из u , которые обладают свойством $P(x)$ [в действительности эту аксиому можно не рассматривать, так как она содержится в аксиоме (6)].

(3) Аксиома пары: $(\forall x)(\forall y)(\exists z)(\forall t)[t \in z \leftrightarrow t = x \vee t = y]$.

Если заданы два множества x и y , то существует множество $\{x, y\}$, элементами которого являются точно x, y .

(4) Аксиома расширенного множества:

$$(\forall x)(\exists y)(\forall z)[z \in y \leftrightarrow z \subset x].$$

Существует множество, элементами которого являются все возможные подмножества заданного множества.

(5) Аксиома пустого множества: $(\forall y)[y \neq \emptyset]$.

Существует множество не имеющее элементов.

(6) Аксиома замены (графическая схема):

$$(\forall x)y(\forall u)(\forall z)[P(x, y) \wedge P(x, u) \rightarrow y = z] \rightarrow (\forall u)(\exists v)(\forall y)[y \in v \leftrightarrow (\exists x)[x \in u \wedge P(x, y)]].$$

Если заданы формула логики предикатов $P(x, y)$ и множество u , то существует множество v , состоящее целиком из множества u , которое соответствует при $P(x, y)$ каждому множеству x [элементы множества обычно являются множествами (4)], принадлежащему u .

(7) Аксиома регулярности:

Любая убывающая последовательность отношений множество — элемент является ограниченной (из того, что элементы множества являются множествами, следует существование убывающей последовательности, продолжающейся как $\dots x_1 x_2 x_3 \dots$. Если эта убывающая последовательность неограниченно продолжается, то возникает противоречие, поэтому она ограничена и должна оканчиваться).

(8) Аксиома суммы множеств: $(\forall x)(\exists y)(\forall z)[z \in y \leftrightarrow (\exists u)[u \in x \wedge z \in u]]$.

Сумма произвольных множеств является множеством.

(9) Аксиома бесконечного множества: $(\exists u)[\varphi \in u \wedge (\forall x)[x \in u \leftrightarrow x \wedge \{x\} \in u]]$.

Существует бесконечное множество.

(10) Аксиома выбора:

$$(\forall z)[(\forall x)[x \in z \rightarrow x \neq \emptyset] \wedge (\forall x)(\forall y)[x, y \in z \wedge x \neq y \rightarrow x \cap y = \emptyset] \rightarrow (\sum u)(\forall x)[x \in z \rightarrow (\exists t)[u \cap x = \{t\}]]].$$

Любое множество имеет функцию выбора. Функция выбора каждому элементу u непустого множества x ставит в соответствие один из элементов u (он обычно является множеством).

Интерпретация аксиом на структуры данных

Исходной посылкой здесь послужит применение множества в качестве основного типа данных при обработке.

(1) Аксиома общности уже используется в данных типа множества в качестве условия равенства двух множеств.

(2) Аксиома включения имеет важнейшее значение, так как определяет эквивалентность понятий, таких, как понятие множества и свойства множества. Здесь множество, является понятием, которое содержит в себе концепцию иерархической структуры; сказанное становится ясным из того, что реальная проблема (например, определение свойств объекта, имеющего заданную структуру, или же наоборот — определение структуры, имеющей требуемые свойства) проявляется в форме получения одного из другого (т.е. из свойств — структуры или наоборот).

(3) Аксиома пары; рекуррентным использованием этой аксиомы порождается упорядоченная группа из n элементов в виде $\langle a_0, a_1, a_2, \dots, a_{n-1} \rangle = \langle a_0, \langle a_1, \langle a_2, \dots, a_{n-1} \rangle \rangle \rangle$.

Это позволяет задать стековую структуру данных.

(4) Аксиома расширенного множества; рекуррентным применением этой теоремы формируется множество с отношением элементов вида $\dots x_1 \in x_2 \in x_3 \dots$. А это не что иное, как иерархическая структура. Конечно, это чисто математическая конструкция, однако в проблемах реального мира иерархия структуры обычно проявляется именно как отношение целое — частная структура. Это совсем не то же самое, что отношение множество — элемент, но создание математической модели как отображение реальной проблемы на мате-

математический мир является важным методом решения проблем, и во многих случаях здесь достаточно присутствия того, что может быть решено путем рассмотрения отношения целое — частная структура через отношение множество — структура элементов.

(5) Аксиома пустого множества; она имеет достаточно общее значение, являясь одним из основных понятий в теории множеств.

(6) Аксиома замены (графическая схема); являясь математическим понятием, она интуитивно трудна для понимания; все множество, которому поставлено в соответствие функциональное отношение, заданное для каждого элемента заданного множества, будет являться тоже множеством или, иными словами, будет расширением части множества, которому поставлено в соответствие исходное множество. Необходимость такой аксиомы следует из того, что, как было сказано в самом начале, требуется определить минимальное множество, которое не порождает противоречий. Прямое произведение является чрезвычайно важной структурой данных в обработке информации. Ее практическая ценность ясна, пожалуй, из того, что основная структура реляционной базы данных описывается прямым произведением.

(7) Аксиома регулярности; ею устанавливается, что иерархическая структура ограничена снизу, т.е. в структуре имеется самый крайний элемент. Иначе говоря, существуют атомы. При реальной обработке на компьютере это является естественным условием.

(8) Аксиома суммы множеств; она входит в аксиому замены (6) и является одной из важнейших операций обработки информации.

(9) Аксиома бесконечного множества не имеет непосредственного отношения к обработке информации.

(10) Аксиома выбора; в математическом смысле — это еще нерешенная проблема, и некоторые математики даже выражают сомнения в ее справедливости. Если же смотреть с позиции технологии обработки информации, то еще большим вопросом, чем корректность аксиомы, является ее практическая реализация. Тем не менее легко можно создать функцию, которая выделяет один элемент из элементов (множеств) множества, а затем уже в соответствии с потребностями определить операции над выбранными элементами.

Базовыми функциями порождения множеств являются следующие: сумма, разность, произведение, прямое произведение, расширение и т.д. Используя эти операции, можно определить новое множество. Например, $*(A \cup B)$ является множеством, построенным из всех подмножеств суммы множеств A и B . Если на основе некоторого уже определенного множества D конструктивно определяется новое множество A , то D принято называть базовым множеством для A . Весьма важным является понятие иерархической структуры, образованной оператором расширения множества. Иерархическая структура определяется рекуррентно как $*n = *(n - 1)$. Язык декларативного типа должен включать все эти основные элементы структуры. Способ представления структуры данных в декларативных языках совершенно отличен от способа ее представления в процедурных методах; здесь на первое место выдвигается способ описания объектов, поскольку в декларативных представлениях семантика символов определяется через непосредственное соответствие между понятиями описываемого мира. В языках процедурного типа «символу» внутри нейрокомпьютера соответствует определенное «действие». Что же касается структуры данных, то если будут «с позиции представления» данных в компьютере подготовлены основные формализмы (не являющиеся основой «семантики»), такие, как массивы и указатели, то можно будет, используя эти формализмы, определить произвольную структуру и можно будет породить самые различные структуры заданием семантики в соответствии с конкретными объектами описания. Все перечисленные обстоятельства привели к многочисленным дискуссиям по теории структуры данных, включающих абстрактные типы данных [5,6].

Формализация представления проблем и ее влияние на нейросетевые технологии

Проблемы, присущие обработке информации на современных нейрокомпьютерах, и те надежды, которые возлагаются на инженерию знаний, позволяют перейти к совершенно новым методам обработки. Целью настоящей статьи является исследование новых методов обработки информации, которые способны решать проблемы в самом широком, насколько это только возможно, диапазоне, и поэтому здесь обсуждаются трудности, присущие традиционной машинной обработке, и те свойства, которых ей недостает. Отметим два типа проблем, отличающихся формализацией описания в сфере обработки информации: детерминированные и недетерминированные. И те, и другие, как это следует из проведенной ранее классификации, наиболее характерны для практической деятельности. Однако, о каком бы из этих типов

ни шла речь, всегда возникает вопрос о достаточных или ограниченных возможностях архитектуры Тьюринга - фон Неймана для их решения. Способы представления проблем этих типов на первый взгляд различны, но по существу они одинаковы. Семантика описания проблем, отличается от семантики, которая может быть воспринята нейрокомпьютером, и основной причиной этого является выполнение семантических преобразований человеком. Особенностью методов описания этих задач, в частности особенностями методов описания человеком, является их сложность. Наиболее, однако, эффективное средство упрощения — это использование человеком языков, включающих непроцедурные представления (семантика декларативного типа), в то время как традиционный компьютер оперирует с языками описания процедурного типа (операционная семантика). Для того, чтобы несколько прояснить ситуацию, обратимся к процессу решения простой проблемы с использованием компьютера с архитектурой Тьюринга - фон Неймана. Рисунок 2 иллюстрирует взаимосвязь реальной проблемы с нейросетью представлением при обработке информации. Этот рисунок состоит, из части (1) — преобразование реальной проблемы в нейрокомпьютере представление после ее предварительной обработки человеком для упорядочения и абстрагирования, части (2) — обработка проблемы внутри нейрокомпьютера и части (3) — представление результатов обработки и их интерпретация. Рисунок показывает, что этот процесс точно соответствует процессу решения проблемы в реальном мире, т.е. результат решения в мире информации должен совпадать с результатом, полученным при использовании тех же самых абстракций или правил преобразования, которые использовались в части (1) для получения результата А. Сплошной линией на рисунке показаны действительно проводимые преобразования, а пунктирной — предполагаемые преобразования.

Рис. 2. Взаимосвязь реальных действий и обработки информации в нейросети:
1 — отображение и интерпретация; 2 — начальный период; 3 — обработка информации;
4 — результаты вычислений.

В этом процессе самая большая нагрузка на оператора, вне всякого сомнения, приходится в части (1). Чтобы пояснить это, обратимся к простому конкретному примеру, представленному на рис. 3. Обычным методом при решении проблем подобного рода является составление модели, описывающей свойства, которыми должен обладать объект исследования, и получение характеристик управления на основе анализа этой модели. Модель представляет собой описание всей информации об объекте, которая необходима для решения поставленной проблемы. Процесс решения показан на рис. 3. Реальная проблема, описанная в крайней левой части рисунка, абстрагируется, и на уровне полученных абстракций составляются модели проблемы. Подчеркнем, что сами модели не имеют отношения к процедурам, которыми будет затем получено решение проблемы, а являются по сути ее описанием. Далее проводится анализ на основе всех этих моделей и осуществляется преобразование к форме, необходимой для процедурного получения решения. Таким образом, можно сказать, что решение проблемы представляет собой процесс последовательного преобразования модельных представлений ее исходной формы вплоть до представления процедурным языком описания. Рисунок 5 иллюстрирует случай, когда уже получено общее решение проблемы, описанной математическими формулами, например путем извлечения некоторой формулы из их множества. Все эти проблемы являются детерминированными проблемами принятия решения, для которых модели полностью заданы и из них могут быть получены атрибуты и поведение объекта исследования. С другой стороны, для недетерминированной проблемы атрибуты и свойства заранее заданы как требования и необходимо получить объект как некоторую сущность, которая им удовлетворит.

Рис. 3. Различные структуры обработки информации и необходимые функции. 1 — воображение, экспрессия, индуктивные выводы, интуиция, структуризация, ассоциация, обучение; 2 — экспрессия, индуктивные выводы, интуиция, структуризация, ассоциация, дедуктивные выводы, понимание, ...; 3 — понимание, ассоциация, структуризация, дедуктивные выводы, ...; 4 — структуризация, обучение, дедуктивные выводы, ассоциация, ...; 5 — вычисление, поиск, декомпозиция, ...; 6 — интуиция, ассоциация, понимание, обучение, дедуктивные выводы, вычисления.

Современные интеллектуальные экспертные системы

Существенная особенность проблемы анализа состоит в том, что процесс ее решения можно разделить на две стадии — составление и исполнение программы, а это все хорошо согласуется со спецификой архитектуры Тьюринга — фон Неймана. Для решения проблем синтеза необходимо прибегнуть к совершенно другому подходу. Обычно в реальном мире существует большое число объектов с одинаковыми функциями, и среди этого многообразия объектов требуется отыскать такой, который обладает всей совокупностью заданных для него функций (см. рис. 4).

Рис. 4. Возможный диапазон компьютерной обработки: 1 — способ представления; 2 — тип проблемы; 3 — синтез; 4 — анализ; 5 — представление на декларативном языке; 6 — представление данных; 7 — представление на процедурном языке; 8 — размер проблемы.

Общий алгоритм решения этой задачи содержит такие действия, как (1) создание исследовательской модели, (2) анализ этой модели как решение проблемы анализа и определение ее функций, (3) сравнение полученных результатов с заданными требованиями и прекращение процесса решения (если результаты и требования совпадают) или же возврат к (2), если совпадение не получено. Этот процесс носит недетер-

минированный, итерационный характер и является более сложным, чем процесс решения проблемы анализа, в том смысле, что включает в свой состав саму проблему анализа. Как следует из рис. 6, пока не существует системы, которая бы эффективно решала сложные проблемы синтеза. Для обеспечения эффективной поддержки необходимо устранить недостатки современных нейросетевых интеллектуальных систем, что вполне осуществимо.

Ниже перечислены особенности которыми должна обладать нейрокомпьютерная система, удовлетворяющая современным требованиям:

- наиболее подходящий тип проблем: анализ, синтез
- возможные формы представления: языковое описание, частично графическое описание;
- масштаб решаемых проблем: большой.

С момента появления нейрокомпьютеров и разделения используемых в ней информационных единиц на данные и команды создалась ситуация, при которой данные пассивны, а команды активны. Все процессы, протекающие в нейрокомпьютере, инициируются командами, а данные используются этими командами лишь в случае необходимости. Для интеллектуальных систем (ИС) эта ситуация неприемлема. Как и у человека, в ИС актуализации тех или иных действий способствуют знания, имеющиеся в системе. Таким образом, выполнение программ в ИС должно инициироваться текущим состоянием информационной базы. Появление в базе фактов или описание событий, установление связей может стать источником активности системы.

Перечисленные основные особенности информационных единиц определяют ту грань, за которой данные превращаются в знания, а базы данных перерастают в базы знаний (БЗ). Совокупность средств, обеспечивающих работу со знаниями, образуют систему управления базой знаний (СУБЗ). В настоящее время не существует баз знаний, в которых в полной мере были бы реализованы перечисленные выше особенности.

Продукционные модели можно считать наиболее распространенными моделями представления знаний. Продукционная модель — это модель, основанная на правилах, позволяющая представить знание в виде предложений типа: «ЕСЛИ условие, ТО действие». Продукционная модель обладает тем недостатком, что при накоплении достаточно большого числа (порядка нескольких сотен) продукций они начинают противоречить друг другу. Системы обработки знаний, использующие продукционную модель получили название «продукционных систем». В состав экспертных систем продукционного типа входят база правил (знаний), рабочая память и интерпретатор правил (решатель), реализующий определенный механизм логического вывода. При разрешении проблемных ситуаций интерпретатор экспертной системы использует данные из базы данных (временные, промежуточные данные) и базы знаний (долгосрочные данные) и формулирует правила, которые приводят к решению задачи.

С целью преодоления сложностей и расширения описательных возможностей логических моделей знаний разрабатываются псевдофизические логики, т.е. логики, оперирующие с нечеткостями, обеспечивающие индуктивные (от частного к общему), дедуктивные (от общего к частному) и традиционные (на одном уровне общности) выводы. Такие расширенные модели, объединяющие возможности логического и лингвистического подходов, принято называть логико-лингвистическими моделями предметной области [7]. Однозначное определение семантической сети в настоящее время отсутствует. В инженерии знаний под ней подразумевается граф, отображающий смысл целостного образа. Узлы графа соответствуют понятиям и объектам, а дуги — отношениям между объектами. Семантическая сеть как модель наиболее часто используется для представления декларативных знаний, т.к. воссоздает модель, схожую с процессами мышления конструктора, технолога. Информация воспринимается специалистом на когнитивном уровне посредством основных единиц этого уровня — понятий, идей, концептов. С помощью этой модели реализуются такие свойства системы знаний, как интерпретируемость и связность. За счет этих свойств семантическая сеть позволяет снизить объем хранимых данных, обеспечивает вывод умозаключений по ассоциативным связям. Одной из первых известных моделей, основанных на семантической сети, является TLC-модель (Teachable Language Comprehender — доступный механизм понимания языка). Модель использовалась для представления семантических отношений между концептами (словами) с целью описания структуры долговременной памяти [8–11].

Как правило, различают экстенциональные и интенциональные семантические сети. Экстенциональная семантическая сеть описывает конкретные отношения данной ситуации. Интенциональная — имена классов объектов, а не индивидуальные имена объектов. Связи в интенциональной сети отражают те отношения, которые всегда присущи объектам данного класса [12–15].

Название понятия (имя фрейма)		
Слоты	Значения	Список связанных процедур
Имя слота 1	Значение 1	Процедура 1
Имя слота 2	Значение 2	Процедура 2.1, Процедура 2.2
...	...	...
Имя слота N	Значение N	Процедура N

Рис. 5. Пример представления знаний фреймовой модели.

Популярным методом представления знаний является и фреймовая модель (Табл. 1). Она представляет собой систематизированную психологическую модель памяти человека и его сознания. Фрейм — это минимально возможное описание сущности какого-либо события, ситуации, процесса или объекта. Существует и другое понимание фрейма — это ассоциативный список атрибутов. Понятие «минимально возможное» означает, что при дальнейшем упрощении описания теряется его полнота, и оно перестает определять ту единицу знаний, для которой было предназначено. Пример представления знаний фреймовой модели представлен на рис. 5.

Представление знаний с помощью фреймов понимается как один из способов представления знаний о ситуациях. Фрейм имеет имя (название) и состоит из слотов. Слоты — это незаполненные (нулевые) позиции фрейма. Если у фрейма все слоты заполнены — это описание конкретной ситуации. В отличие от моделей других типов во фреймовых моделях фиксируется жесткая структура информационных единиц, которая называется протофреймом. Значением слота может быть практически что угодно (числа, математические соотношения, тексты на естественном языке или на языке программ, ссылки на другие слоты данного фрейма). Значением слота может выступать и отдельный фрейм, что является очень удобным для упорядочивания знаний по степени общности. Исключение из фрейма любого слота делает его неполным, а иногда и бессмысленным. При конкретизации фрейма ему и слотам приписываются конкретные имена и происходит заполнение слотов. Таким образом из протофреймов получаются фреймы — экземпляры. Переход от исходного протофрейма к фрейму — экземпляру может быть многошаговым, за счет постепенного уточнения значений слотов [16].

Математическая структура данных в декларативных языках

Декларативные языки как языки представления знаний составляют ядро систем обработки знаний. Однако декларативные языки помимо рассмотренных уже условий должны удовлетворять множеству других, характерных именно для языков представления знаний, поэтому проектирование языка должно вестись на достаточно строгой математической основе.

Уже говорилось, что в языке декларативного типа существуют отношения непосредственного соответствия между объектами описываемого мира, их отношения между собой, между атрибутами и символическими последовательностями, создаваемыми на основе грамматических правил языка. Другими словами, этот язык состоит из части, которая соответствует (1) объектам «внутри описываемого мира», и части, которая соответствует (2) отношениям и атрибутам (функциональные отношения также входят в состав обычных отношений). Внутри компьютера первая часть представляет структуру данных, а вторая — относящиеся к ним описания. Далее, при обсуждении форм представления в качестве особенностей декларативного языка была рассмотрена возможность представления совершенно однородной формой описания, например предикатными формулами, любых структурных отношений внутри описываемых объектов. Однородную систему описания (не содержащую структуры внутри элементов описания) принято называть плоским представлением. В таком случае о языке декларативного типа можно сказать, что он является языком, принимающим плоскую языковую форму. Но одновременно с этим декларативный язык можно преобразовать к языку, не являющемуся плоским, посредством смысловой замены структуре данных части предикатов. Поскольку синтаксис плоской формы весьма прост, то она является крайне привлекательной с точки зрения теоретического изящества, однако это проявляется в основном при использовании плоского языка в качестве языка представления знаний для сравнительно небольших проблем. Для больших про-

блем, наиболее важных с практической стороны, язык с плоской формой весьма беден, как о том был сделан вывод в предыдущем разделе.

В нашем естественном языке содержится понятие структуры сущностей в самых разнообразных формах, более того, в представлениях несколько более сложных проблем непременно используются графические средства [17–20]. В этой теории исследуется формальное представление объектов исходя из посылки использования предикатов для описания отношений. Далее такие сущности (множества), как прямые произведения, которые непосредственно связаны с (структурными) отношениями, могут также применяться для описания самых широких понятий. Все перечисленное выходит за рамки данных здесь определений и задается как некое множество, не содержащее в себе противоречий в том смысле, что не приводит к возникновению каких-либо парадоксов.

Заключение

Таким образом, каждой аксиоме аксиоматической теории множеств можно поставить в соответствие систему обработки информации, которая осуществляет задание свойств множества, определение множества (структура данных) и определение операций над множеством. Важным обстоятельством является наличие нескольких базовых функций, в частности функций, относящихся к генерации новых множеств. Если, положим, в системе обработки информации некоторые из свойств отсутствуют, то применение системы там, где эти функции необходимы, будет невозможным. Традиционных возможностей файловых систем оказывается недостаточно для построения даже простых информационных нейросетей. В работе выявлено несколько потребностей, которые не покрываются возможностями систем управления базами знаний: поддержание логически согласованного набора файлов; обеспечение языка манипулирования данными; восстановление информации после разного рода сбоев; реально параллельная работа нескольких пользователей при использовании нейрокомпьютеров. Предполагается, что если прикладная информационная система опирается на некоторую систему управления технологическими данными, обладающую этими свойствами, то она является системой управления базами данных. Тем не менее, полезно сформулировать эти (и другие) важные функции в виде модели представления знаний. В статье предложены аппаратные модели создания базы данных аппаратов и технологических процессов. Каждая из них обладает своими плюсами и минусами, и поэтому для каждой конкретной задачи необходимо выбрать именно свою модель. Уникальность предложенного алгоритма состоит в том, что они позволяют автоматизировать получение логических выводов на основании вновь поступающих фактов и формализованных экспертных знаний. Знания могут быть как непосредственно введены или импортированы в систему, так и получены при обработке любой доступной информации, в т.ч. с использованием нейросетей/машинного обучения. Поддержка принятия решений необходима в различных областях науки и техники и сферах деятельности, где требуются обоснованные, логически доказуемые аргументы для принятия решений — в оперативном и стратегическом управлении. Для принятия стратегического решения необходимо проведение детального моделирования последствий предполагаемого решения, поиск оптимального пути достижения заданного результата при помощи технологий имитационного моделирования.

Список литературы

1. Сигов А. С., Нечаев В. В., Кошкарёв М. И. Архитектура предметно-ориентированной базы знаний интеллектуальной системы // *International Journal of Open Information Technologies*. 2014. № 12. С. 1–6.
2. Гаврилова Г. А., Хорошевский В. Ф. Базы знаний интеллектуальных систем. СПб.: Питер. 2000. 384 с.
3. Джаррантано Д., Райли Г. Экспертные системы: принципы разработки и программирование. 4-е изд.: Пер. с англ. М.: Вильямс, 2007. 1152 с.: ил. Парал. тит. англ. ISBN: 978-5-8459-1156-8.
4. Helbig, H. *Knowledge Representation and the Semantics of Natural Language*/ Springer. Berlin, Heidelberg, New York. 2006. P. 655. ISBN: 978-3-540-24461-5.
5. Сидняев Н. И. Нейросети и нейроматематика: учебное пособие/ Н.И. Сидняев, П.В. Храпов: под ред. Н.И. Сидняева. М.: Издательство МГТУ им. Н.Э. Баумана. 2016. 83 с.
6. Сидняев Н. И., Бутенко Ю. И., Болотова Е. Е. Экспертная система производственного типа для создания базы знаний о конструкциях летательных аппаратов // *Аэрокосмическое приборостроение*. 2019. № 6. С. 38–52. DOI: 10.25791/Aviakosmos.06.2019.676
7. Крислов В. А., Побережник С. М., Тарасенко Р. А. Сравнительный анализ моделей представления знаний в интеллектуальных системах // *Тр. Одес. политехн. ун-та, Одесса*. 1998. № 2. С. 45–49.
8. Дошина А. Д. Экспертная система. Классификация. Обзор существующих экспертных систем // *Молодой ученый*. 2016. № 21. С. 756–758.
9. Логунова Е. А. Обзор подходов к разрешению недостатков производственной базы знаний системы логического вывода // *Современные наукоемкие технологии*. 2015. № 9. С. 46–48.
10. Елисеев Д. В. Модель представления знаний при создании адаптивной информационной системы // *Наука и образование: научное издание МГТУ им. Н.Э. Баумана*. 2010. № 03. С. 1–6.

11. Белоус Е. С., Кудинов В. А., Желнин М. Э. Современные модели представления знаний в обучающих системах // Ученые записки. Электронный научный журнал Курского государственного университета. 2010. № 1 (13). С. 9–14.
12. Хабаров С. П. Интеллектуальные информационные системы. PROLOG-язык разработки интеллектуальных и экспертных систем: учеб. пособ. // СПб.: СПбГЛТУ. 2013. С.138.
13. Братко И. Алгоритмы искусственного интеллекта на языке PROLOG, 3-е издание.: Пер. с англ. М.: Издательский дом «Вильямс». 2004. 640 с. ISBN: 5-8459-0664-4
14. Гаврилова Т. А., Кудрявцев Д. В., Муромцев Д. И. Инженерия знаний. Модели и методы. М.: «Лань», 2016. 324 с. ISBN: 978-5-8114-2128-2
15. Хрусталева Е. Ю. Логико-лингвистические модели наукоемкого производственного комплекса как разновидность интеллектуальных информационных систем // Экономический анализ: теория и практика. 2014. № 11 (363). С. 11–22.
16. Овчинева Ю. А. Семантическая сеть — перспективная платформа для системы управления знаниями // Вестник ГУУ. 2015. № 3. С. 14–16.
17. Соколова О. В. Категория фрейма в когнитивной лингвистике // Вестник АГТУ. 2007. № 1. С. 236–239.
18. Вахталова Ю. С. Место признака в структуре фреймовых моделей представления знания // Вестник СПбГУ: Язык и литература. 2007. № 1 (ч. 2). С. 232–237.
19. Минский М. Фреймы для представления знаний. М.: Энергия, 1979. 152 с.
20. Bruynooghe, M. et al. Predicate Logic as a Modeling Language: Modeling and Solving Some Machine Learning and Data Mining Problems with IDP3 // Theory and Practice of Logic Programming. 2015. Т. 15. No. 6. P. 783–817. DOI: <https://doi.org/10.1017/s147106841400009x>

Поступила 12 июня 2020 г.

PREDICATE LOGIC IN INFORMATION PROCESSING SYSTEMS IN KNOWLEDGE BASES

E. E. BOLOTOVA, YU. I. BUTENKO, AND N. I. SIDNYAEV

doi: 10.25210/jfop-2002-037047

The article reveals the basic principles of building intelligent systems for knowledge engineering. Research on artificial intelligence using neurosystems and methods of knowledge representation in expert systems is presented. The logic of predicates, syntax and semantics of predicate language, rules of inference in predicate logic, as well as ways of presenting information in computer systems at various stages of problem solving, models of solving and classification of problems are discussed. The basic structures of knowledge processing systems and knowledge representation languages are considered. The structural and functional scheme of the intellectual system for the neurocomputer implementation of constructive operational-advising expert systems is developed. It is shown that the core of any intellectual system is a knowledge base that describes a specific subject area. It is postulated that the choice of the correct model of knowledge in expert systems plays an important role, and in most cases becomes the basis for choosing an approach to solving a particular problem.

НОВЫЕ МАТЕМАТИЧЕСКИЕ И ФИЗИЧЕСКИЕ МЕТОДЫ

УДК 517.51+517.518.85

НОВОЕ СЕМЕЙСТВО БЕСКОНЕЧНО ГЛАДКИХ ФУНКЦИЙ С КОМПАКТНЫМ НОСИТЕЛЕМ

© Автор, 2020

doi: 10.25210/jfor-2002-048059

Коновалов Я. Ю. — ст. преподаватель, Московский государственный технический университет им. Н. Э. Баумана, Москва. E-mail: Kon2000200@mail.ru

Аннотация

В работе исследованы свойства нового семейства бесконечно гладких функций $\text{fip}_{a,n}(x)$ с компактным носителем, представляющих собой свертки В-сплайнов Шенберга и атомарных функций $h_a(x)$. Данное семейство включает в себя семейство атомарных функций $\text{fup}_n(x)$ и является его естественным обобщением. При целых значениях параметра a и $a = \sqrt[n]{n}$ функции $\text{fip}_{a,n}(x)$ являются атомарными. Представлены формулы для вычисления предложенных функций, рассмотрена методика их применения при интерполяции и решении краевых задач. Описаны преимущества и недостатки новых функций по сравнению с В-сплайнами и семейством $\text{fup}_n(x)$. Приведены результаты решения модельных задач.

Ключевые слова: сплайны, атомарные функции, свертка, интерполяция, краевая задача

Abstract

In the work properties of new infinitely smooth compactly supported functions $\text{fip}_{a,n}(x)$ are considered. These functions are defined as convolutions of atomic functions $h_a(x)$ and B-splines. This new family includes family of atomic functions $\text{fup}_n(x)$ and presents its native generalization. If a is integer or $a = \sqrt[n]{n}$, then $\text{fip}_{a,n}(x)$ are atomic functions. Formulae for computation of proposed functions are presented. Method of application new functions to interpolation and boundary-value problems is considered. Advantages and disadvantages of new functions in comparison to B-splines and $\text{fup}_n(x)$ family are described. Results of solution of model problems are demonstrated.

Keywords: splines, atomic functions, convolution, interpolation, boundary-value problem

Введение

При решении прикладных задач широко используется локальный базис В-сплайнов Шенберга [1]. Этот базис обеспечивает наименьшую погрешность при данной ширине ленты матрицы. Кроме того, В-сплайны, представляющие собой кусочно-полиномиальные функции, легко вычислять. Известен и недостаток этого базиса. В-сплайн порядка n имеет не более чем $n - 1$ непрерывную производную, что затрудняет применение сплайнов при решении дифференциальных уравнений и интерполяции функций высокой степени гладкости. Одним из путей преодоления этой проблемы является использование функций, близких по форме к сплайнам, но обладающих большей гладкостью. Примером подобных функций являются, в частности, атомарные функции. Среди них наиболее близким к сплайнам является семейство функций $\text{fup}_n(x)$, представляющих собой свертки атомарной функции $\text{up}(x)$ с В-сплайнами. В данной работе предлагается новое семейство функций, являющихся свертками В-сплайнов и атомарных функций $h_a(x)$, являющихся обобщением функции $\text{up}(x)$.

1. В-сплайны

По определению В-сплайн n -го порядка Θ_n представляет собой кратную свертку $n + 1$ прямоугольного импульса $\varphi(x)$

$$\Theta_n(x) = \underbrace{\varphi(x) * \varphi(x) * \dots * \varphi(x)}_{n+1},$$

$$\text{где } \varphi(x) = \begin{cases} 1, & |x| < \frac{1}{2}, \\ \frac{1}{2}, & |x| = \frac{1}{2}, \\ 0, & |x| > \frac{1}{2}. \end{cases}$$

В-сплайн Θ_n имеет носитель $\left[-\frac{n+1}{2}; \frac{n+1}{2}\right]$.

Преобразования Фурье прямоугольного импульса и В-сплайна равны соответственно

$$\hat{\varphi}(t) = \text{sinc}\left(\frac{t}{2}\right) \text{ и } \widehat{\Theta}_n(t) = \text{sinc}^{n+1}\left(\frac{t}{2}\right). \quad (1)$$

Непосредственно из формулы (1) для спектра $\Theta_n(x)$ вытекает тождество

$$\Theta'_n(x) = \Theta_{n-1}\left(x + \frac{1}{2}\right) - \Theta_{n-1}\left(x - \frac{1}{2}\right). \quad (2)$$

В-сплайны Θ_n обеспечивают разложение единицы

$$\sum_{k \in \mathbb{Z}} \Theta_n(x - k) \equiv 1$$

и точное представление многочленов степени не выше n

$$\sum_{k \in \mathbb{Z}} c_k \Theta_n(x - k) \equiv P_m(x), \quad m \leq n.$$

Известно [1], что при приближении сдвигами $\Theta_n(x)$ на сетке с шагом h обеспечивается погрешность порядка h^{n+1} при ширине ленты матрицы n . Наиболее широко используется кубический В-сплайн $\Theta_3(x)$, обеспечивающий 4-й порядок погрешности и позволяющий представить систему уравнений для неизвестных коэффициентов разложения c_k в виде трехдиагональной матрицы.

2. Атомарные функции

Атомарные функции [2] – это финитные (с компактным носителем) решения дифференциальных уравнений с линейно преобразованным аргументом вида

$$L(y) = \sum_{k=1}^K c_k y(ax - b_k), \quad (3)$$

где $L(y)$ — дифференциальный оператор, $a > 1$, b_k и c_k — постоянные.

2.1. Семейство атомарных функций $h_a(x)$

Для любого вещественного $a > 1$ атомарные функции $h_a(x)$ [2–5] являются решениями уравнения

$$y'(x) = \frac{a^2}{2} (y(ax + 1) - y(ax - 1)) \quad (4)$$

и удовлетворяют условию нормировки

$$\int_{-\frac{1}{a-1}}^{\frac{1}{a-1}} h_a(x) dx = 1. \quad (5)$$

Носителем функции $h_a(x)$ является отрезок $\left[-\frac{1}{a-1}; \frac{1}{a-1}\right]$. При $a > 2$ функция $h_a(x)$ принимает постоянное значение $h_a(x) = \frac{a}{2}$ на отрезке $\left[-\frac{a-2}{a(a-1)}; \frac{a-2}{a(a-1)}\right]$. С ростом a относительная длина участка постоянства увеличивается. Таким образом, $h_a(x)$ можно рассматривать как сглаженный прямоугольный импульс. Сдвиги $h_a(x)$ на $\frac{2}{a}$ образуют разложение единицы

$$\sum_{k \in \mathbb{Z}} h_a\left(x - \frac{2}{a}k\right) \equiv \frac{a}{2}. \quad (6)$$

В то же время $h_a(x)$ сама является сверткой прямоугольных импульсов убывающей длины и имеет спектр

$$\widehat{h}_a(t) = \prod_{k=1}^{\infty} \text{sinc}\left(\frac{t}{a^k}\right). \quad (7)$$

Свертки $h_a(x)$ с самой собой (атомарные функции $\text{ch}_{a,n}(x)$) соответственно являются своего рода атомарными аналогами сплайнов. Их свойства подробно рассмотрены в работах [6–9].

Моменты атомарной функции h_a являются рациональными числами. Известны итерационные формулы для их точного вычисления [3]. Так как $h_a(x)$ — четная функция, ее нечетные симметричные моменты равны нулю.

Для четных моментов

$$a_{2n} = \int_{-\frac{1}{a-1}}^{\frac{1}{a-1}} x^{2n} h_a(x) dx$$

имеем

$$a_{2n} = \frac{(2n)!}{a^{2n} - 1} \sum_{k=0}^{n-1} \frac{a_{2k}}{(2k)!(2n-2k+1)!}. \quad (8)$$

В силу (5) $a_0 = 1$. Для несимметричных нечетных моментов

$$b_{2n-1} = \int_0^{\frac{1}{a-1}} x^{2n-1} h_a(x) dx$$

имеем

$$b_{2n-1} = \frac{1}{a^{2n-1} 4n} \sum_{k=0}^n a_{2k} C_{2n}^{2k}.$$

При $a = 2$ $h_a(x)$ совпадает с атомарной функцией $up(x)$, имеющей носитель $[-1; 1]$.

2.2. Семейство атомарных функций $fup_n(x)$

Атомарные функции $fup_n(x)$ [3,4] определяются как свертки $up(x)$ и В-сплайна порядка $n-1$

$$fup_n(x) = up(x) * \Theta_{n-1}(x).$$

Соответственно, $fup_0(x) \equiv up(x)$.

Они являются решениями функционально-дифференциальных уравнений вида

$$y'(x) = \frac{1}{2^{n-1}} \sum_{k=0}^{n+2} (C_{n+1}^k - C_{n+1}^{k-1}) y\left(2x + \frac{n+2}{2} - k\right).$$

Аналогично В-сплайнам

$$fup'_n(x) = fup_{n-1}\left(x + \frac{1}{2}\right) - fup_{n-1}\left(x - \frac{1}{2}\right). \quad (9)$$

Длина носителя $fup_n(x)$ равна сумме длин носителей $up(x)$ и $\Theta_{n-1}(x)$ и составляет $n+2$.

Интерполяционные свойства базиса $fup_n(x-k)$ легко оценить из следующих соображений. Свертка $up(x)$ с многочленом дает многочлен той же степени. Тогда

$$\sum_{k=-\infty}^{+\infty} c_k fup_n(x-k) = \sum_{k=-\infty}^{+\infty} c_k (\Theta_{n-1}(t) * up(t)) \Big|_{x-k} = up(x) * \sum_{k=-\infty}^{+\infty} c_k \Theta_n(x-k) \equiv up(x) * P_{n-1}(x) \equiv Q_{n-1}(x),$$

где $P_{n-1}(x)$ и $Q_{n-1}(x)$ — некоторые многочлены степени $n-1$. Таким образом, сумма сдвигов $fup_n(x)$ обеспечивает точное приближение лишь многочленов степени $n-1$, обеспечивая погрешность порядка h^n . Кроме того, так как длина носителя $fup_n(x)$ равна $n+2$, трехдиагональная матрица системы позволяет использовать лишь базис $fup_2(x)$ с погрешностью $O(h^2)$. Для получения погрешности порядка h^4 требуется базис $fup_4(x)$ и пятидиагональная матрица. Несмотря на этот недостаток, атомарные функции $fup_n(x)$ нашли широкое применение [10,11].

2.3. Свертки атомарных функций

Приведем формулировку теоремы о свертке двух атомарных функций.

Пусть $f(x)$ и $g(x)$ — две атомарные функции, удовлетворяющие уравнениям (10) и (11) с одинаковым масштабным коэффициентом a

$$f^{(n_f)} = \sum_{m=1}^M c_{f_m} f(ax - b_{f_m}), \quad (10)$$

$$g^{(n_g)} = \sum_{k=1}^K c_{g_k} g(ax - b_{g_k}). \quad (11)$$

Тогда их свертка $h(x) = f(x) * g(x)$ также является атомарной функцией и удовлетворяет уравнению

$$h^{(n_f+n_g)} = \sum_m \sum_k \frac{c_{f_m} c_{g_k}}{a} h(ax - b_{f_m} - b_{g_k}).$$

Доказательство теоремы приведено в [6,8], более общая формулировка в [12].

Отметим, что для порядков производных в левых частях уравнений (10) и (11) n_f и n_g допускаются не только положительные значения, но и равенство нулю как одного из них, так и обоих.

Это позволяет применять приведенную выше теорему в случае, когда одна из свертываемых функций является прямоугольным импульсом, в частности, при выводе функциональных уравнений для В-сплайнов и функционально-дифференциальных уравнений для семейства $\text{fir}_n(x)$.

3. Новое семейство бесконечно гладких финитных функций $\text{fir}_{a,n}(x)$

3.1. Определение

Для любого вещественного $a > 1$ любого $n \geq 1$ определим функцию $\text{fir}_{a,n}(x)$ как свертку

$$\text{fir}_{a,n}(x) = h_a(x) * \Theta_{n-1}(x). \quad (12)$$

При $n=0$ функция $\text{fir}_{a,0}(x) = h_a(x)$. Из определения (12) следует

$$\text{fir}_{2,n}(x) = \text{fir}_n(x), \text{fir}_{a,0}(x) = h_a(x), \text{fir}_{2,0}(x) = \text{ur}(x), \text{fir}_{a,n}(x) * \Theta_m(x) = \text{fir}_{a,n+m}(x).$$

Длина носителя $\text{fir}_{a,n}(x)$ равна сумме длин носителей $h_a(x)$ и $\Theta_{n-1}(x)$, следовательно, в отличие от $\text{fir}_n(x)$, к n прибавляется не 2, а сколь угодно малое $\frac{2}{a-1}$. Таким образом, для трехдиагональной матрицы при $n=3$ и $a \geq 3$ можно получить погрешность порядка h^3 , что лучше, чем для $\text{fir}_2(x)$, хотя и хуже, чем для В-сплайна. При больших значениях параметра a возможен переход к квазиинтерполяции с $n=4$.

3.2. Основные свойства

Рассмотрим некоторые свойства $\text{fir}_{a,n}(x)$.

Спектр $\text{fir}_{a,n}(x)$ равен произведению спектров $h_a(x)$ (7) и В-сплайна (1)

$$\widehat{\text{fir}_{a,n}}(t) = \text{sinc}^n\left(\frac{t}{2}\right) \prod_{k=1}^{\infty} \text{sinc}\left(\frac{t}{a^k}\right). \quad (13)$$

Из выражения (13) для спектра аналогично (2) и (9) вытекает

$$\text{fir}'_{a,n}(x) = \text{fir}_{a,n-1}\left(x + \frac{1}{2}\right) - \text{fir}_{a,n-1}\left(x - \frac{1}{2}\right). \quad (14)$$

Доказательство: из вида преобразования Фурье (13) следует

$$\widehat{\text{fir}_{a,n}}(t) = \widehat{\text{fir}_{a,n-1}}(t) \cdot \text{sinc}\left(\frac{t}{2}\right) = \widehat{\text{fir}_{a,n-1}}(t) \cdot \frac{e^{\frac{it}{2}} - e^{-\frac{it}{2}}}{(t/2) \cdot 2i}.$$

Домножим полученное выражение на it

$$it \cdot \widehat{\text{fir}_{a,n}}(t) = \widehat{\text{fir}_{a,n-1}}(t) \cdot \left(e^{\frac{it}{2}} - e^{-\frac{it}{2}}\right).$$

Обратное преобразование Фурье последнего равенства дает требуемое

$$\text{fir}'_{a,n}(x) = \text{fir}_{a,n-1}\left(x + \frac{1}{2}\right) - \text{fir}_{a,n-1}\left(x - \frac{1}{2}\right).$$

Дифференцируя (14) и подставляя в правую часть (14), получим

$$\text{fir}''_{a,n} = \text{fir}_{a,n-2}(x+1) - 2\text{fir}_{a,n-2}(x) + \text{fir}_{a,n-2}(x-1).$$

Далее аналогичным путем получим для производной порядка $m \leq n$

$$\text{fir}_{a,n}^{(m)} = \sum_{k=0}^m C_m^k (-1)^k \text{fir}_{a,n-m}\left(x + \frac{m}{2} - k\right). \quad (15)$$

При $m=n$ (15) обращается в

$$\text{fir}_{a,n}^{(n)} = \sum_{k=0}^n C_n^k (-1)^k h_a\left(x + \frac{n}{2} - k\right). \quad (16)$$

Вычисление производных $\text{fir}_{a,n}(x)$ порядка выше n должно осуществляться путем дифференцирования слагаемых (16) по формуле (4).

Сдвиги функции $\text{fir}_{a,n}(x)$ на 1 дают разложение единицы и точное представление многочленов вплоть до степени $n-1$

$$\sum_{k \in \mathbb{Z}} \text{fir}_{a,n}(x-k) \equiv 1,$$

$$\sum_{k \in \mathbb{Z}} c_k \text{fir}_{a,n}(x-k) \equiv P_m(x), \quad m \leq n-1. \quad (17)$$

Доказательство

$$\sum_{k \in \mathbb{Z}} \text{fir}_{a,n}(x-k) = \sum_{k \in \mathbb{Z}} \int_{-\infty}^{\infty} h_a(t) \Theta_{n-1}(x-k-t) dt = \int_{-\infty}^{\infty} h_a(t) \sum_{k \in \mathbb{Z}} \Theta_{n-1}(x-k-t) dt$$

Известно, что В-сплайны дают разложение единицы, следовательно

$$\int_{-\infty}^{\infty} h_a(t) \sum_{k \in \mathbb{Z}} \Theta_{n-1}(x-k-t) dt = \int_{-\infty}^{\infty} h_a(t) dt = 1.$$

Что и требовалось доказать. Иллюстрация приведена на рис. 1.

Точное представление многочленов доказывается аналогично.

Рис. 1. Разложение единицы сдвигами $\text{fir}_{3,3}$ на 1.

Интересно, что из (6) следует, что сдвиги функции $\text{fir}_{a,n}(x)$ на $\frac{2}{a}$ также дают разложение единицы

$$\sum_{k \in \mathbb{Z}} \text{fir}_{a,n}\left(x - \frac{2}{a}k\right) = \sum_{k \in \mathbb{Z}} \int_{-\infty}^{\infty} \Theta_{n-1}(t) h_a\left(x - \frac{2}{a}k - t\right) dt =$$

$$= \int_{-\infty}^{\infty} \Theta_{n-1}(t) \sum_{k \in \mathbb{Z}} h_a\left(x - \frac{2}{a}k - t\right) dt = \int_{-\infty}^{\infty} \Theta_{n-1}(t) \cdot \frac{a}{2} dt = \frac{a}{2},$$

однако, точного представления полиномов они не обеспечивают. Иллюстрация приведена на рис. 2.

Рис. 2. Разложение единицы сдвигами $\text{fir}_{3,3}$ на $2/3$.

Для вычисления $\text{fir}_{a,n}(x)$ можно использовать ряд Фурье

$$\text{fir}_{a,n}(x) = \frac{2}{l} \left(\frac{1}{2} + \sum_{k=1}^{\infty} \widehat{\text{fir}_{a,n}} \left(\frac{2\pi k}{l} \right) \cos \left(\frac{2\pi kx}{l} \right) \right), \text{ где } l = n + \frac{2}{a-1}.$$

3.3. Функции $\text{fir}_{a,n}(x)$ при натуральных значениях параметра a

При натуральных значениях параметра a функции $\text{fir}_{a,n}(x)$ атомарные. Построим для них уравнения вида (3). Заметим, что $\forall m \in \mathbb{N}$ прямоугольный импульс $\varphi(x)$ удовлетворяет уравнению

$$y(x) = \sum_{k=1}^m y \left(mx - k + \frac{m+1}{2} \right). \quad (18)$$

Тогда, согласно теореме о свертке атомарных функций [6,8,12], может быть построено уравнение для свертки $h_m(x) * \varphi(x) = \text{fir}_{m,1}(x)$, затем и для свертки $\text{fir}_{m,1}(x) * \varphi(x) = \text{fir}_{m,2}(x)$ и т.д. Таким образом, при целых значениях a , $\text{fir}_{a,n}(x)$ являются атомарными функциями и удовлетворяют функционально-дифференциальным уравнениям первого порядка. Хотя записать явный вид уравнений для произвольных a и n проблематично, в каждом конкретном случае его легко получить путем последовательного применения теоремы. Найденные уравнения можно использовать для построения итерационных алгоритмов [8].

Выведем уравнение для $\text{fir}_{3,3}(x)$.

Запишем уравнение для прямоугольного импульса

$$y(x) = y(3x+1) + y(3x) + y(3x-1)$$

и уравнение для $\text{fir}_{3,0}(x) = h_3(x)$

$$y'(x) = \frac{9}{2} (y(3x+1) - y(3x-1)).$$

Осуществим свертку. Получим уравнение для $\text{fir}_{3,1}(x)$

$$y'(x) = \frac{3}{2} (y(3x+2) + y(3x+1) - y(3x-1) - y(3x-2)).$$

Далее получим уравнение для $\text{fir}_{3,2}(x) = \text{fir}_{3,1}(x) * \varphi(x)$

$$y'(x) = \frac{1}{2} (y(3x+3) + 2y(3x+2) + 2y(3x+1) - 2y(3x-1) - 2y(3x-2) - y(3x-3))$$

и, наконец, уравнение для

$$\text{fir}_{3,3}(x) = \text{fir}_{3,2}(x) * \varphi(x),$$

$$y'(x) = \frac{1}{6} (y(3x+4) + 3y(3x+3) + 5y(3x+2) + 4y(3x+1) - 4y(3x-1) - 5y(3x-2) - 3y(3x-3) - y(3x-4)). \quad (19)$$

3.4. Функции $\text{fir}_{a,n}(k)$ при $a = \sqrt[k]{m}$

Для применения теоремы о свертке атомарных функций требуется равенство масштабных коэффициентов в уравнениях свертываемых функций. Следовательно, для свертки с прямоугольным импульсом, заданным уравнением (18) требуется функция, заданная уравнением с целым масштабным коэффициентом. Построим такие уравнения для функций $h_a(x)$ при $a = \sqrt[k]{m}$.

Возьмем производную от обеих частей уравнения (4) для $h_a(x)$

$$y''(x) = \frac{a^3}{2} (y'(ax+1) - y'(ax-1)).$$

Подставим в правую часть выражение (4). Получим для $h_a(x)$ уравнение второго порядка с масштабным коэффициентом a^2

$$y''(x) = \frac{a^5}{4} (y(a^2x+a+1) - y(a^2x+a-1) - y(a^2x-a+1) + y(a^2x-a-1)).$$

При $a = \sqrt{m}$, где $m > 1$ - натуральное число, этот масштабный коэффициент будет равен m . Таким образом, становится возможным при помощи теоремы о свертке атомарных функций построить уравнение для свертки $h_{\sqrt{m}}(x) * \varphi(x) = \text{fir}_{\sqrt{m},1}(x)$. Далее можно последовательно построить уравнения для свертки $\text{fir}_{\sqrt{m},n}(x) = \text{fir}_{\sqrt{m},n-1}(x) * \varphi(x)$.

Например, при $a = \sqrt{3}$ получим для $h_{\sqrt{3}}(x)$ уравнение 2-го порядка

$$y'' = \frac{9}{4}\sqrt{3}(y(3x + \sqrt{3} + 1) - y(3x + \sqrt{3} - 1) - y(3x - \sqrt{3} + 1) + y(3x - \sqrt{3} - 1)).$$

Тогда уравнение для $\text{fir}_{\sqrt{3},1}(x)$ имеет вид

$$\begin{aligned} y'' = & \frac{3}{4}\sqrt{3}(y(3x + \sqrt{3} + 2) + y(3x + \sqrt{3} + 1) - \\ & - y(3x + \sqrt{3} - 1) - y(3x + \sqrt{3} - 2) - \\ & - y(3x - \sqrt{3} + 2) - y(3x - \sqrt{3} + 1) + \\ & + y(3x - \sqrt{3} - 1) + y(3x - \sqrt{3} - 2)), \end{aligned}$$

а для $\text{fir}_{\sqrt{3},2}(x)$

$$\begin{aligned} y'' = & \frac{\sqrt{3}}{4}(y(3x + \sqrt{3} + 3) + 2y(3x + \sqrt{3} + 2) + 2y(3x + \sqrt{3} + 1) - \\ & - 2y(3x + \sqrt{3} - 1) - 2y(3x + \sqrt{3} - 2) - y(3x + \sqrt{3} - 3) - \\ & - y(3x - \sqrt{3} + 3) - 2y(3x - \sqrt{3} + 2) - 2y(3x - \sqrt{3} + 1) + \\ & + 2y(3x - \sqrt{3} - 1) + 2y(3x - \sqrt{3} - 2) + y(3x - \sqrt{3} - 3)). \end{aligned}$$

Далее, повторяя процедуру свертки, можно получить уравнения второго порядка для $\text{fir}_{\sqrt{3},n}(x)$ при сколь угодно большом n .

Общий вид уравнения порядка k для $h_a(x)$ может быть получен индукцией по k .

$$y^{(k)} = \left(\frac{a^2}{2}\right)^k \cdot a^{\frac{k(k-1)}{2}} \cdot \sum_{s=0}^{2^k-1} (-1)^{\sum_{l=0}^{k-1} p_l} y\left(a^k x + \sum_{l=0}^{k-1} (-1)^{p_l} a^l\right). \quad (20)$$

Каждое из слагаемых правой части (20) имеет аргумент $a^k x$, к которому прибавлена сумма степеней a от 1 до a^{k-1} , взятых со всеми возможными комбинациями знаков. При этом, если число минусов в аргументе четно, то слагаемое берется с плюсом, если нечетно, то с минусом. Для удобства перебора знаков слагаемых можно интерпретировать их как разряды двоичной записи целого числа s , изменяющегося от 0 до $2^k - 1$. Уравнение (4) является частным случаем (20) при $k=1$ и может рассматриваться как база индукции.

При подстановке $a = \sqrt[k]{m}$ в (20) получим уравнение с целым масштабным коэффициентом. Далее при помощи теоремы о свертке атомарных функций для любых натуральных $m > 1$ и k можно последовательно получить уравнения для $\text{fir}_{\sqrt[k]{m},1}(x)$, $\text{fir}_{\sqrt[k]{m},2}(x)$ и далее $\text{fir}_{\sqrt[k]{m},n}(x)$ для сколь угодно большого n . Уравнения будут иметь порядок k . Следовательно, функции $\text{fir}_{\sqrt[k]{m},n}(x)$ также являются атомарными.

3.5. Вычисление значений $\text{fir}_{a,n}(x)$

Для решения прикладных задач полезно знать точные значения функции $\text{fir}_{a,n}(x)$ в определенных точках. Например, для интерполяции наиболее важны значения функции в узлах интерполяции, соответствующих целым значениям аргумента. Используя явное выражение $\Theta_n(x)$ в виде кусочно-полиномиальной функции, можно свести вычисление сверточного интеграла (12) к моментам атомарной функции $h_a(x)$.

Вычислим значение $\text{fir}_{3,3}(0) = \frac{17}{24}$

$$\text{fir}_{3,3}(0) = \Theta_2(t) * h_3(t) \Big|_{x=0} = \int_{-\infty}^{\infty} \Theta_2(t) h_3(t-x) dt \Big|_{x=0} = \int_{-\frac{1}{2}}^{\frac{1}{2}} \Theta_2(t) h_3(t) dt.$$

Последнее верно т.к. $\text{supp}(h_3) = \left[-\frac{1}{2}; \frac{1}{2}\right]$. Так как на этом промежутке $\Theta_2 = \frac{3}{4} - t^2$

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} \Theta_2(t) h_3(t) dt = \int_{-\frac{1}{2}}^{\frac{1}{2}} \left(\frac{3}{4} - t^2\right) h_3(t) dt = \frac{3}{4} - a_2 = \frac{17}{24},$$

где a_2 второй момент h_3 , вычисляемый по формуле для моментов (8) и равный $\frac{1}{24}$.

Значение $\text{fir}_{3,3}(1) = \frac{7}{48}$

$$\begin{aligned} \text{fir}_{3,3}(1) &= \Theta_2(t) * h_3(t) \Big|_{x=1} = \int_{-\infty}^{\infty} \Theta_2(t) h_3(t-x) dt \Big|_{x=1} = \int_{\frac{1}{2}}^{\frac{3}{2}} \Theta_2(t) h_3(t-1) dt = \\ &= \int_{\frac{1}{2}}^{\frac{3}{2}} \frac{1}{2} \left(\frac{3}{2} - t\right)^2 h_3(t-1) dt = \Big| t = y+1 \Big| = \frac{1}{2} \int_{-\frac{1}{2}}^{\frac{1}{2}} \left(\frac{1}{2} - y\right)^2 h_3(y) dy = \\ &= \frac{1}{2} \int_{-\frac{1}{2}}^{\frac{1}{2}} \left(\frac{1}{4} - y + y^2\right) h_3(y) dy = \frac{1}{2} \left(\frac{1}{4} + a_2\right) = \frac{1}{2} \left(\frac{1}{4} + \frac{1}{24}\right) = \frac{7}{48}. \end{aligned}$$

Используя полученное ранее уравнение для $\text{fir}_{3,3}(x)$, найдем значения $\text{fir}'_{3,3}(-1)$.

Подставим в (19) $x = -1$. Получим

$$\begin{aligned} \text{fir}'_{3,3}(-1) &= \frac{1}{6} (\text{fir}_{3,3}(1) + 3\text{fir}_{3,3}(0) + 5\text{fir}_{3,3}(-1)) = \\ &= \frac{1}{6} \left(\frac{7}{48} + 3 \frac{17}{24} + 5 \frac{7}{48} \right) = \frac{1}{6} \cdot \frac{144}{48} = \frac{1}{2}. \end{aligned}$$

Продифференцируем (19)

$$\begin{aligned} y''(x) &= \frac{1}{2} (y'(3x+4) + 3y'(3x+3) + 5y'(3x+2) + 4y'(3x+1) - \\ &- 4y'(3x-1) - 5y'(3x-2) - 3y'(3x-3) - y'(3x-4)). \end{aligned}$$

Рис. 3. Графики $\text{fir}_{3,3}(x)$ и ее производных.

Подставим $x = 0$ и $x = 1$. Получим, соответственно

$$\text{fir}''_{3,3}(0) = 2\text{fir}'_{3,3}(1) - 2\text{fir}'_{3,3}(-1) = -2, \quad \text{fir}''_{3,3}(1) = \frac{1}{2} (-5\text{fir}'_{3,3}(1) - 3\text{fir}'_{3,3}(0) - \text{fir}'_{3,3}(-1)) = 1.$$

Найденные значения $\text{fir}_{3,3}(x)$ и ее производных приведены в табл. 1.

Графики функции $\text{fir}_{3,3}(x)$ и ее первой и второй производной приведены на рис. 3.

Таблица 1. Точные значения $\text{fir}_{3,3}(x)$ и ее производных

$x =$	-1	0	1
$\text{fir}_{3,3}(x) =$	$7/48$	$17/24$	$7/48$
$\text{fir}'_{3,3}(x) =$	$1/2$	0	$-1/2$
$\text{fir}''_{3,3}(x) =$	1	-2	1

4. Некоторые применения $\text{fir}_{a,n}(x)$

4.1. Интерполяция сдвигами $\text{fir}_{a,n}(x)$

Сдвиги на 1 $\text{fir}_{a,n}(x)$ дают интерполяционную формулу точную для полиномов степени $n-1$

$$S(x) = \sum_k c_k \text{fir}_{a,n}(x-k). \quad (21)$$

Носитель $\text{fir}_n(x)$ имеет длину $l(a,n) = n + \frac{2}{a-1}$. Для того, чтобы при интерполяции по формуле (21) система уравнений имела трехдиагональную матрицу требуется $2 < l(a,n) \leq 4$. Носитель такой длины получается

$$\text{для } n=1 \text{ при } \frac{5}{3} \leq a < 3,$$

$$\text{для } n=2 \text{ при } a \geq 2,$$

$$\text{для } n=3 \text{ при } a \geq 3.$$

Рассмотрим интерполяцию сдвигами $\text{fir}_{a,3}(x)$. В силу вышесказанного она точна для многочленов 2-й степени. При $a \geq 3$ коэффициенты c_k удовлетворяют системе с трехдиагональной матрицей вида

$$\begin{pmatrix} \text{fir}_{a,3}(1) & \text{fir}_{a,3}(0) & \text{fir}_{a,3}(-1) & & & y_0 \\ & \text{fir}_{a,3}(1) & \text{fir}_{a,3}(0) & \text{fir}_{a,3}(-1) & & y_1 \\ & & \ddots & \ddots & \ddots & \vdots \\ & & & \text{fir}_{a,3}(1) & \text{fir}_{a,3}(0) & \text{fir}_{a,3}(-1) & y_n \end{pmatrix}.$$

Дополним систему условием на первые производные $S(x)$ на концах

$$\begin{aligned} c_{-1} \text{fir}'_{a,3}(1) + c_1 \text{fir}'_{a,3}(-1) &= y'_0 \\ c_{n-1} \text{fir}'_{a,3}(1) + c_{n+1} \text{fir}'_{a,3}(-1) &= y'_n. \end{aligned} \quad (22)$$

Полученная система уравнений может быть решена при помощи итерационных алгоритмов.

При шаге сетки h отличном от 1 необходимо в правой части (22) умножить $y'(x_i)$ на h .

4.2 Решение краевых задач для дифференциальных уравнений

Выражение (17) обеспечивает равномерное приближение для производных

$$S^{(l)}(x) = \sum_{k \in \mathbb{Z}} c_k \text{fir}_{a,n}^{(l)}(x+k).$$

Это позволяет использовать $\text{fir}_{a,n}(x)$ для решения дифференциальных уравнений. Рассмотрим уравнение $p(x)y'' + q(x)y' + r(x)y = f(x)$ с условиями на концах $y(x_0) = y_0$, $y(x_1) = y_1$. Система с трехдиагональной матрицей будет содержать строки вида

$$\begin{aligned} & \left(p(x_0 + ih) \frac{1}{h^2} \text{fir}_{a,n}''(1) + q(x_0 + ih) \frac{1}{h} \text{fir}_{a,n}'(1) + r(x_0 + ih) \text{fir}_{a,n}(1) \right) c_{i-1} + \\ & + \left(p(x_0 + ih) \frac{1}{h^2} \text{fir}_{a,n}''(0) + r(x_0 + ih) \text{fir}_{a,n}(0) \right) c_i + \\ & \left(p(x_0 + ih) \frac{1}{h^2} \text{fir}_{a,n}''(-1) + q(x_0 + ih) \frac{1}{h} \text{fir}_{a,n}'(-1) + r(x_0 + ih) \text{fir}_{a,n}(-1) \right) c_{i+1} = f(x_0 + ih). \end{aligned}$$

Краевые условия представлены двумя дополнительными уравнениями

$$\begin{aligned} c_{-1} \text{fir}_{a,n}(1) + c_0 \text{fir}_{a,n}(0) + c_1 \text{fir}_{a,n}(-1) &= y(x_0), \\ c_{n-1} \text{fir}_{a,n}(1) + c_n \text{fir}_{a,n}(0) + c_{n+1} \text{fir}_{a,n}(-1) &= y(x_n). \end{aligned}$$

Аналогично [9,13] рассмотрим модельные задачи вида $-y'' + r(x)y = f(x)$ с условиями $y(0) = 0$, $y(1) = 0$ и точным решением $\bar{y}(x)$. Явный вид $r(x)$, $f(x)$ и $\bar{y}(x)$ приведен в табл. 2.

Таблица 2. Условия модельных краевых задач

№	$r(x)$	$f(x)$	$\bar{y}(x)$
1	x	$(4\pi^2 + x)\sin 2\pi x$	$\sin 2\pi x$
2	e^x	$(\pi^2 + e^x)\sin \pi x$	$\sin \pi x$
3	$(x - 0.1\pi)^{-1}$	$(\pi^2 + r(x))\sin \pi x$	$\sin \pi x$
4	$2\text{tg}^2 x + 1$	$-(1 + 2\text{tg}^2 x) \cdot \left(\frac{x}{\cos 1} + 1 - x \right)$	$\frac{1}{\cos x} - \frac{x}{\cos 1} + x - 1$
5	$\sqrt{x}$	$-\frac{15}{4}\sqrt{x} + x^3 - \sqrt{x^3}$	$\sqrt{x^5} - x$

Выберем шаг сетки $h = 0.05$ как в [9,13]. Сравним результаты, полученные при использовании базиса $\text{fir}_{a,3}(x)$ с результатами, полученными ранее с использованием В-сплайнов и атомарных функций $\text{fur}_2(x)$ [13] и $\text{ch}_{a,n}(x)$ [9]. Графики базисных функций $\Theta_3(x)$, $\text{fur}_2(x)$ и $\text{fir}_{3,3}(x)$ приведены на рис 4.

Рис. 4. Графики базисных функций $\Theta_3(x)$, $\text{fur}_2(x)$ и $\text{fir}_{3,3}(x)$.

Из табл. 3 видно, что во всех примерах $\text{fir}_{a,3}(x)$ обеспечивают меньшую погрешность чем $\text{fur}_2(x)$. Атомарные функции $\text{ch}_{a,n}(x)$ удовлетворяют функционально-дифференциальному уравнению порядка n ($\text{ch}_{3,n}(x)$ — 3-го порядка), следовательно, производные $\text{ch}'_{a,n}(x)$ и $\text{ch}''_{a,n}$ нужно отдельно вычислять, например как суммы ряда Фурье. В то же время производные функций $\text{fir}_{a,n}(x)$ явно выражаются по формулам (14), (15), (16), а при целых a $\text{fir}_{a,n}(x)$ удовлетворяет функционально-дифференциальному уравнению первого порядка. Кроме того, возможно точное вычисление значений $\text{fir}_{a,n}(x)$ в узлах коллокации. Поэтому, хотя базис $\text{ch}_{a,n}(x)$ в некоторых примерах обеспечивает меньшую погрешность, применение $\text{fir}_{a,n}(x)$ в задачах коллокации более удобно и требует меньших вычислительных затрат.

Таблица 3. Результаты решения модельных краевых задач

№	Базис	$\Delta_{\max}$	$y_{\max}$	$x_{\max}$
1*	Θ_3	0.00845	0.75	-1
1*	fup_2	0.00559	0.75	-1
1**	$ch_{5,3}$	0.00461	0.275	0.983
1**	$ch_{10,3}$	0.00423	0.278	0.980
2*	Θ_3	0.00175	0.5	1
2*	fup_2	0.00164	0.5	1
2**	$ch_{5,3}$	0.00099	0.473	0.995
2**	$ch_{10,3}$	0.00091	0.471	0.995
3*	Θ_3	0.00282	0.45	0.985
3*	fup_2	0.00216	0.45	0.985
3**	$ch_{5,3}$	0.00172	0.309	0.824
3**	$ch_{10,3}$	0.00149	0.310	0.826
4*	Θ_3	0.00084	0.7	-0.288
4*	fup_2	0.00056	0.7	-0.288
4**	$ch_{5,3}$	0.00048	0.735	-0.278
4**	$ch_{10,3}$	0.00045	0.735	-0.278
5*	Θ_3	0.00023	0.2	-0.182
5*	fup_2	0.00017	0.2	-0.182
5**	$ch_{5,3}$	0.00015	0.164	-0.153
5**	$ch_{10,3}$	0.00014	0.165	-0.153

№	Базис	$\Delta_{\max}$	$y_{\max}$	$x_{\max}$
1	$fip_{3,3}$	0.00546	0.27	0.99
1	$fip_{4,3}$	0.00545	0.27	0.99
1	$fip_{10,3}$	0.00545	0.27	0.99
1	$fip_{20,3}$	0.00568	0.27	0.99
2	$fip_{3,3}$	0.00117	0.478	0.998
2	$fip_{4,3}$	0.00119	0.478	0.997
2	$fip_{10,3}$	0.00121	0.479	0.998
2	$fip_{20,3}$	0.00142	0.481	0.998
3	$fip_{3,3}$	0.00204	0.306	0.82
3	$fip_{4,3}$	0.00187	0.308	0.824
3	$fip_{10,3}$	0.00174	0.310	0.827
3	$fip_{20,3}$	0.00205	0.310	0.827
4	$fip_{3,3}$	0.00069	0.737	-0.277
4	$fip_{4,3}$	0.00052	0.687	-0.291
4	$fip_{10,3}$	0.00035	0.637	-0.298
4	$fip_{20,3}$	0.00036	0.999	-0.002
5	$fip_{3,3}$	0.00016	0.163	-0.152
5	$fip_{4,3}$	0.00017	0.213	-0.192
5	$fip_{10,3}$	0.0002	0.312	-0.258
5	$fip_{20,3}$	0.00026	0.462	-0.317

Результаты, отмеченные * взяты из [13], ** — из [9].

Выводы

Предложено новое двухпараметрическое семейство бесконечно гладких базисных функций $fip_{a,n}(x)$. При целых значениях параметра a , а также при $a = \sqrt[k]{m}$ функции $fip_{a,n}(x)$ являются атомарными. Численный эксперимент показывает перспективность применения новых функций в методах интерполяции и коллокации. Некоторые результаты данной работы были представлены на международных конференциях [14,15].

Список литературы

- Shikin, E.V., Plis, A. I. Handbook on Splines for the User. CRC Press, 1995. ISBN 0-8493-9404-X
- Рвачев В.Л., Рвачев В.А. Неклассические методы теории приближений в краевых задачах. Киев: Наукова думка 1979.
- Кравченко В.Ф. Лекции по теории атомарных функций и некоторым их приложениям. Монография. М.: Радиотехника, 2003. ISBN: 5-93108-019-8
- Кравченко В. Ф., Кравченко О. В. Конструктивные методы алгебры логики, атомарных функций, вейвлетов, фракталов в задачах физики и техники. М.: Техносфера, 2018. ISBN: 978-5-94836-518-3
- Кравченко В.Ф., Чуриков Д.В. Цифровая обработка сигналов атомарными функциями и вейвлетами. М.: Техносфера, 2019. (доп. тираж)
- Кравченко В.Ф., Коновалов Я.Ю., Пустовойт В.И. Новый класс окон на основе семейства атомарных функций $Ch_{a,n}(x)$ и его применение в цифровой обработке сигналов // Радиотехника и электроника, 2015, Т. 60. № 9. С. 931-943. DOI: 10.7868/S0033849415090065
- Кравченко В.Ф., Коновалов Я.Ю., Пустовойт В.И. Семейства атомарных функций $Ch_{a,n}(x)$ и $Fup_n(x)$ в цифровой обработке сигналов // Доклады академии наук. 2015. Т. 462. № 1. С. 35-40. DOI: 10.7868/S0869565215130083
- Kononov, Y.Y. Iterative Algorithms for Computation Convolutions of Atomic Functions Including New Family $Ch_{a,n}$. Days on Diffraction 2012. International Conference. Saint Petersburg, May 28 – June 1, 2012. Proceedings. P. 129-133. DOI: 10.1109/DD.2012.6402765
- Kononov, Y.Y., Kravchenko, O.V. Application of New Family of Atomic Functions $Ch_{a,n}$ to Solution of Boundary Value Problems. Proceedings of the International Conference «Days on Diffraction 2014», St. Petersburg, 2014. P. 132-137. DOI: 10.1109/DD.2014.7036438

10. Gotovac, H., Cvetkovic, V., and Andricevic, R. Adaptive Fup Multi-Resolution Approach to Flow and Advective Transport in Highly Heterogeneous Porous Media: Methodology, Accuracy and Convergence. *Advances in Water Resources*. Vol. 32 (2009). No. 6. P. 885-905.
DOI: 10.1016/j.advwatres.2009.02.013
11. Rvachova, T.V., Tomilova, Ye. P. Finding Antiderivatives with the Help of the Generalized Taylor Series // *Открытые информационные и компьютерные интегрированные технологии*. № 73. 2016. P. 52-58.
12. Рвачев В.А. Атомарные функции. Математический анализ и теория вероятностей. Сборник научных трудов. К.: Наукова думка. 1978. С. 143-146.
13. Рвачев В.Л., Федотова Е.А. Сопоставление аппроксимационных свойств сплайнов и атомарных функций. Методы сплайн-функций // *Вычислительные системы*, 72. Сборник трудов. Новосибирск, 1977. С. 92-98.
14. Коновалов Я.Ю. О некоторых свертках атомарных функций и В-сплайнов Шенберга. Акустооптические и радиолокационные методы измерений и обработки информации: Материалы 11-й Международной научно-технической конференции // *Российское НТОРЭС им. А.С. Попова*. Суздаль. Россия. 2018. С. 23-26.
15. Konovalov, Y.Y. New Infinitely Differentiable Spline-Like Basis Functions // 2019 Photonics and Electromagnetics Research Symposium - Spring (PIERS-Spring), Rome, Italy, 2019. P. 114-122. DOI: 10.1109/PIERS-Spring46901.2019.9017707

Поступила 21 апреля 2020 г.

NEW FAMILY OF INFINITELY SMOOTH COMPACTLY SUPPORTED FUNCTIONS

YA.YU KONOVALOV

doi: 10.25210/jfop-2002-048059

The work is devoted to the construction of a new family of infinitely smooth compactly supported functions $\text{fip}_{a,n}(x)$. Their properties are investigated. A numerical experiment to solve boundary value problems for linear differential equations is carried out.

Functions $\text{fip}_{a,n}(x)$ are defined as convolutions of atomic function $h_a(x)$ and B spline of $n-1$ degree. They are infinitely smooth like $h_a(x)$ and have approximative properties similar to B splines. Arbitrary polynomial of $n-1$ degree may be exactly represented as sum of integer shifts of $\text{fip}_{a,n}(x)$.

Derivatives of order m $\text{fip}_{a,n}^{(m)}(x)$ are represented in explicit form by sums of shifts of $\text{fip}_{a,n-m}(x)$ and $h_a(x)$ if $m \geq n$. If a is integer then $\text{fip}_{a,n}(x)$ satisfies functional-differential equation of the first order. Algorithm of construction of such equation basing on the atomic functions convolution theorem is described. If a equals to the root of the integer number then $\text{fip}_{a,n}(x)$ satisfies functional-differential equations of order equal to the degree of root. In this case to allow application of convolution of atomic functions theorem special equations of higher order are constructed for $h_a(x)$. In both cases $\text{fip}_{a,n}(x)$ are atomic functions. For example functional-differential equations for $\text{fip}_{3,n}(x)$ if $n=1,2,3$ and $\text{fip}_{\sqrt{3},n}(x)$ if $n=1$ and $n=2$ are constructed.

Integral representing convolution of $h_a(x)$ and B spline may be expressed by the moments of $h_a(x)$ and then computed exactly. That allows computation of exact values of $\text{fip}_{a,n}(x)$ and their derivatives. Exact values of $\text{fip}_{3,3}(x)$ and its first and second derivatives are found.

Method of application of $\text{fip}_{a,n}(x)$ functions to interpolation and collocation problems is described. To check efficiency of the method comparison of results of model boundary value problems solution basing on new function and known bases of B splines, atomic functions $\text{fup}_2(x)$ and family of atomic functions $\text{ch}_{a,n}(x)$ is performed. Numerical experiment demonstrates advantage of new functions in comparison to of B splines, and $\text{fup}_2(x)$. Approximation quality of $\text{fip}_{a,n}(x)$ basis is similar to $\text{ch}_{a,n}(x)$. Main advantage of $\text{fip}_{a,n}(x)$ compared to $\text{ch}_{a,n}(x)$ is more convenient application and less computation cost. In case of $\text{ch}_{a,n}(x)$ derivatives of less order, especially first and second, must be computed separately due to $\text{ch}_{a,n}(x)$ are solutions of equations of order n . This requires additional computations and increases error. In other hand there are explicit expressions for each derivative of $\text{fip}_{a,n}(x)$ and values of $\text{fip}_{a,n}(x)$ in grid points may be found exactly.

OPTICAL MATERIALS AND TECHNOLOGIES

UDC 778.534.19

GLASSLESS 3D IMAGING SYSTEMS BASED ON PROJECTORS AND MOBILE PHONES

© Authors, 2020

doi: 10.25210/jfop-2002-060073

Petrov N. I. is a chief researcher, Dr. Sci. (Phys.-Math), Scientific and Technological Centre of Unique Instrumentation of RAS. E-mail: petrovni@mail.ru

Sokolov Yu. M. is a head of Laboratory, senior scientist, PhD (Phys.-Math), Peoples Friendship University of Russia (RUDN University). Email: sokolov_yuri@mail.ru.

Khromov M. N. is a senior scientist, PhD (Phys.-Math), VNIIFTRI, Mendeleevo, Moscow region. E-mail: mak.hrom@mail15.com

Аннотация

Разработана многовидовая система 3D-отображения изображений, включающая модули съемки трехмерных объектов, обработки изображений (создания файлов 3D-изображений) и оптический экран для воспроизводства 3D-изображений на основе интегральной технологии визуализации. Созданы большие экраны дисплея с использованием набора проекторов. Мульти-проекторная система используется для улучшения характеристик дисплея, таких как разрешение, глубина и угол обзора. Разработаны также 3D дисплеи на основе мобильных устройств. Файлы 3D-изображений созданы с использованием файлов глубины и 2D-изображений, снятых под разными углами обзора. Продемонстрированы 3D изображения с использованием набора Full HD проекторов и проекторов Ultra HD 4K. Показана возможность наблюдения скрытых объектов, расположенных за передним элементом, с использованием проекторов и смартфонов 4K Sony (разрешение 3840×2160), а также Samsung и LG (разрешение 2560×1440).

Ключевые слова: 3D-дисплей, 3D-изображение, интегральное изображение, многопроекторный дисплей, мобильный 3D-дисплей, создание файла 3D-изображения

Abstract

A multi-view 3D display system which includes the modules of capturing three-dimensional objects, image processing (creation of 3D image files) and display screen for the 3D image displaying based on integral imaging technology has been developed. Large display screens using multi-projectors are created. Multi-projector system is used for improving the performance, such as viewing resolution, depth and viewing angle. 3D displays based on mobile devices are also developed. 3D image files are created using 2D + depth files and 2D images captured at different view angles. Large 3D images using Full HD multi-projectors and Ultra HD 4K projector are created. Images having the property of "looking around" an object are demonstrated using projectors and smartphones 4K Sony (resolution 3840×2160), Samsung and LG (resolution 2560×1440).

Keywords: 3D display, 3D image, integral imaging, multi-projector display, mobile 3D display, 3D image file creation

Introduction

There are different types of the real 3D display systems, including multi-view, holographic and integral imaging methods [1–3]. Integral imaging 3D display is one of the promising displays that provide different perspectives according to viewing direction. Integral imaging is the most simple and suitable method of creating real 3D image. It has essential advantages in comparison with other methods (absence of the complicated mechanical and optical system, use of inexpensive materials in manufacture, possibility of simultaneous vision by the several viewers and others), which allow producing autostereoscopic 3D display with satisfactory color image. However, there are still some physical limitations. Limitations of current displays are small view angle, the small image depth and the presence of sharp edges at the transition between adjacent viewing zones.

The resolution and the image size can be increased if multi-projection system is used. Different types of projection integral imaging systems have been proposed for various applications [4–8]. Autostereoscopic displays having a large screen, for example, more than 100 inches are generally constructed using multiple projectors. A multi-view display system that combines multiple flat-panel 3D displays through a multi-projection system was proposed in order to increase the number of views [9]. The multi-view display module using the four projectors was proposed to realize a 14.4-inch large-screen 3D display in [10]. A 100-inch 300-Mpixel multi-projection 3D display using 300 projectors was developed in [11]. Multi-view autostereoscopic projection display, using a pair of micro-electromechanical (MEMS) scanner based projectors, a head-tracking camera, and a rotating retro-reflective diffuser screen is demonstrated in [12]. In [13] a resolution enhanced integral imaging display method using two micro-lens

arrays (MLA) with different focal lengths for capturing and display respectively is proposed. A computational multi-projection display is proposed by employing a multi-projection system combining with compressive light field displays in [14]. In [15, 16] preliminary design of the 3D system based on integral imaging was carried out. The real-time integral imaging scheme was realized and experimentally demonstrated. A method that enlarges the viewing zone of an integral 3D image using multiple projectors was proposed in [17]. Multi-projectors are used for improving the performance, such as viewing resolution, viewing angle, etc. It is of practical interest to transfer these 3D techniques to the smartphones. In [18] a conventional commercial mobile 3D display system is analyzed with simulations and experiments. The mobile phone with a diamond pentile display panel was used to reduce a crosstalk. In [19] an optical 3D integral imaging display with a smartphone using elemental images from the quick-response codes was demonstrated.

In this paper the multi-view 3D display system which includes the modules of capturing three-dimensional objects, image processing (creation of 3D image files) and display screen for the 3D image displaying based on integral imaging technology is developed. Large display screens using multi-projectors are created. Multi-projector system is used for improving the performance, such as viewing resolution, viewing angle, etc. 3D display based on mobile devices is also analyzed. Images on smartphones 4K Sony (resolution 3840×2160) and Samsung and LG (resolution 2560×1440) are demonstrated.

1. Multi-view projector-based display system

The design of the 3D system includes several steps: capture of 3D images, creating 3D image files and displaying 3D images.

1.1. 3D Image capturing

The integral imaging method and capturing object from different directions using a rotating platform and a conventional camera were used. The drawback of the integral imaging method for capturing is the small depth. Capturing object from different angles with a camera allows the higher depth to be obtained (Fig. 1). 2D images were captured using rotating table at different view angles in 1–2 degrees intervals. Note that the capturing method using a camera and a rotating table allowed us to apply simple mathematical model to create 3D image files (see Section 1.2).

Fig. 1. Rotating table for image capturing by a camera.

1.2. Creation of 3D image files

To create a 3D image, the original 2D images are used, the number of which corresponds to the number of views. We place a portion of each view under each lens of the lenticular sheet to be able to see each view at a strictly defined angle (Fig. 2). In this case, the order of views in the neighboring lenses is the same. In the first stage, the original views are converted to images with a vertical resolution equal to the vertical resolution of the screen and with a

horizontal resolution equal to the number of lenses involved in forming three-dimensional images. Each view is then divided into vertical strips one pixel thick. We place one strip from each view under one vertical cylindrical lens, i.e. the vertical columns of pixels are aligned with the lenticular lenses. A column of pixels is assigned to one view for each of the N columns. Then the image is formed under the next vertical lens and so on until the end of the formation of the entire image. The result is an image with a horizontal resolution of NM , where N is the number of views and M is the resolution of one view, usually equal to the number of lenses in the lenticular involved in the 3D image. A similar procedure is used to create 3D files for four projectors. Codes for integration of images captured at different angles into one common 3D file are created in MatLab.

For optimal image quality, the system is adjusted so that under each lens there are an integer number of pixels, each of which is responsible for the formation of a separate view of 3D image (Fig. 2b).

Fig. 2. Schematic representation of the formation of 3D images using an array of lenticulars.

The basic design of the system consisting of a projector, diffuser layer and lenticular lens is presented in Fig. 3. As shown in Fig. 3, the light rays from a display element of the projector are incident on a screen, which produces viewpoints for observes.

Fig. 3. Schematic representation of the 3D display system. 1 — projector; 2 — display element; 3 — projection lens; 4 — elemental image; 5 — diffuser layer; 6 — lenticular lens; 7 — viewing zone; 8 — observer; φ — viewing angle.

The viewing angle of the integral imaging display is defined by the lens-pitch of the lenticular sheet and the focal length of the elemental lens

$$\varphi = 2\arctan\left(\frac{D}{2f}\right) = 2\arctan\left(\frac{D}{2(h-R)}\right), \quad (1)$$

where D is the lens-pitch of the lens-array, f is the focal length and R is the curvature radius of the elemental lens in the lenticular sheet, $h = f + R$ is the total thickness of the lenticular sheet.

Fig. 4. Multi-projector 3D imaging system: (a) front view; (b) rear view of 53-inch screen and 4 HD projectors.

1.3. Experimental system

An experimental display system was designed and constructed (Fig. 4). Four projectors with a 4K equivalent resolution are arranged in a 2×2 array horizontally and vertically. The display screen consists of a diffuser layer scattering uniformly in both vertical and horizontal directions and lenticular sheet. Here we present a simple case of the vertical orientation of the lenticular lenses. In [15] we considered also the case of slanted axes of the lenticular lenses. The specifications of the prototype are presented in Table 1. The projectors and screen are adjusted on the optical table with a simple pitch test method. The multi-projector system (Fig. 4) is constructed to increase the size and resolution of the 3D image. Four full HD digital light processing (DLP) Acer projectors are combined together. The problems of alignment of boundaries, colors, etc. are solved. The boundaries between images and color and brightness variations from different projectors are eliminated by optical and image processing methods (see Section 2.2). A 53-inch size screen (Fig. 4) is designed and 3D images are demonstrated.

Gap distance between the projectors and screen depends on the pitch size of the lenticular lens and number of viewpoints. This distance is equal to 50 cm for the lenticular sheet with the pitch $D = 1.27$ mm (20 lpi) and number of viewpoints $N = 6$.

Table 1. Parameters of the display prototype

Lenticular array (vertically aligned)	Pitch	1.27 mm (20 lpi)
	Focal length	2.04 mm
	Thickness	3.25 mm
Projector	Resolution	1920×1080 pixels
	Number of units	4
3D image	Resolution	4K/view numbers
	Size	744 mm (H)×420 mm (V)
	Viewing angle	35°

1.4. Combining images from four projectors

Four projectors, the images of which are slightly overlapped as shown in Fig. 5, are used in order to increase the resolution and image size. Several problems arise in this case: precise alignment between the projected images, light brightness, and color matching among all projectors, etc.

Fig. 5. Layout of images on the screen when using 4 projectors.

To compensate for the brightness increase at the intersections of the images, the images were corrected (Fig. 6). It can be seen that the brightness in the overlapping parts is higher than in other parts (Fig. 5). Images were corrected by the image processing to compensate for the increased brightness at the intersection of the images of adjacent projectors. The brightness of the entire image can be smoothed by reducing the brightness of the pixels corresponding to the overlapping parts. Along each axis, the brightness of the image changes over a part of the image (set in the MatLab program) according to the expression x^{-a} , where x is the pixels coordinate and a is the power index. The optimal power index depends on the optics of the projectors and it was found experimentally. In our case, it was found, that $a = 3/2$.

Fig. 6. Brightness adjustment of an image from a single projector.

1.5. Alignment of projectors and screen

Precise alignment between the projected images and screen is required in order to obtain high quality 3D image.

The distance between the screen and projector is determined by the parameters of the projector, lenticular sheet, and view numbers. The size of the 3D image depends on the number of view N for the given lenticular sheet and display matrix of the projector. The image size in horizontal direction is defined as $L = (m/N)D$, where m is the resolution (pixel number in horizontal direction) of the projector's display matrix, N is the view number, and D is the pitch of the lenticular sheet. We obtain that the image size $L = 406$ mm for the number of views $N = 6$ and $L = 271$ mm when $N = 9$ for the lenticular sheet with 20 lpi and matrix with $m = 1920$.

Precise alignment between the projectors and the screen can be achieved with a pitch test. The pitch test is a very important step in 3D imaging. The need for the pitch test is to avoid misalignment between the lenticules on the lenticular sheet and projected image. To adjust the position of the screen relative to the projector, a special image is used, which is a set of strips of two colors (in most cases black and white). Black and white stripes are visible on the screen (Fig. 7a).

When aligning, it is necessary to obtain the most uniform image of the same color, which indicates that each lens of the lenticular array contains the exact number of pixels corresponding to the number of views in the 3D image (Fig. 7b). The presence of two color bars in the 3D image shows the difference between the width of the lenticular lens and the part of the image corresponding to one lens. This means that the part of the image that should be under one lens is wider or narrower than the lens itself, so «beats» occur between the structure of the lenticular sheet and the structure of the image. Projection distortions due to the misalignment of the projector and the lenticular screen occur when the projectors are not exactly in the horizontal plane. The alignment procedure can be significantly accelerated with a mirror in front of the screen. First, the screen should be installed on the optical table. Then the spatial position of the projector can be smoothly changed according to the change of the pitch test pattern on the mirror.

Fig. 7. Pitch-test images on the screen. (a) misalignment due to incorrect distance between the lenticular lens sheet and projector; (b) correct alignment between the lenticular lenses and projected image.

2. 3D Image Displaying: Experimental results

In Fig. 8 the images of 15-inch size from different viewing directions are shown. Here we used the BENQ full HD DLP projector (resolution 1920×1080 pixels). 3D image has a “look-around” capability. Objects, which are hidden for the view in one view angle, become visible at another viewing angle. 3D images with 6–8 views were created using own developed software program in MatLab. Lenticular sheets with different thicknesses and diffuser layers were used in the experiments.

Fig. 8. 3D images from different viewing directions. Arrows show the relative positions of different parts of fighters.

Image dimensions can be increased while maintaining the same resolution by increasing the number of projectors. To increase the depth the lenticular arrays with aspheric surface profiles should be developed. Multi-projector system allows the large-screen 3D display to be created. In Fig. 9 the images from different viewing directions on

53-inch size screen are shown. The image size is 34-inches: 744 mm (H)×420 mm (V). Here we used four full HD DLP Acer projectors (resolution 1920×1080 pixels). 3D images with 6–12 views were created. Lenticular sheets with 20 lpi were used in the experiments. 3D image has a “look-around” capability. Objects (LED source placed behind the fighter’s foot), that are hidden for the view in the bottom image become visible at another viewing angle (the upper image). There is no viewing distance limitation: continuous images can be viewed at distances $z > 1$ m. The viewing zone at the center has a viewing angle of 35° and shows 3D images with smooth motion parallax at distances from 1 m to over 5 m. The value of the depth corresponds to the real depth of the captured 3D object. There are three viewing zones (the primary viewing zone and two secondary viewing zones). Discontinuity occurs only at the transition from the primary viewing zone to the secondary viewing zone.

It should be noted that there are still some problems: low resolution, depth and view angle. It is followed from the experiments that the greater the viewing angle, the less the depth range. To increase the depth the lenticular arrays with aspheric surface profiles should be developed. Two-dimensional microlens arrays should be designed in order to obtain motion parallax in vertical direction. Crosstalk (ghosting) causes a major perceptual problem in the 3D display system and it is a critical factor determining the image quality of autostereoscopic displays [20, 21]. However, as has been shown in [22], crosstalk in the multiview 3D images is not an effective parameter of defining the quality of 3D images.

Fig. 9. 34-inch images on 53-inch screen. Images at different observation angles:
(a) the picture shows a red LED, which is located behind the foot of the warrior (b).

In Fig. 11 images using Ultra HD4K projector (Fig. 10) with resolution 3840×2160 pixels are presented. Unlike the use of four HD projectors, here we don’t have to solve problems of alignment of boundaries and colors.

Fig. 10. Ultra HD4K projector.

Fig. 11. 40 inch 3D images using 4K ultra HD projector.

3. Mobile 3D Display

Mobile 3D display consists of a commercial mobile phone 2D display and a lenticular lens array. We used smartphones with different pixel structures (Fig. 12): 4K Sony (resolution 3840×2160) and Samsung and LG (resolution 2560×1440). Specifications are presented in Table 2.

Table 2. Specifications of mobile phones

Mobile device	Screen size, in	Resolution		Display matrix type
		Pixels number	Pixel size, μm	
LG G3	5.52	2560×1440 (534 ppi)	46	IPS LCD
4K Sony Xperia XZ Premium	5.46	3840×2160 (807 ppi)	29	IPS LCD
Samsung Galaxy Note 4	5.7	2560×1440 (515 ppi)	49	OLED Diamond pixel
Samsung Galaxy S7	5.1	2560×1440 (577 ppi)	44	OLED Diamond pixel

Fig. 12. Images of pixel structures: (a) LG G3 — IPS; (b) 4K Sony Xperia — IPS;

(c) Samsung Galaxy Note 4 — diamond pixel structure;

(d) Samsung Galaxy S7 — diamond pixel structure. Numbers from 1 to 6 corresponds to the view numbers.

3.1. Creation of 3D files

2D+z (2D+depth) image files from available resources were also used as the basis for creating 3D images. A three-dimensional model of the object was created based on the depth and color data at each point (Fig. 13). Peculiar 2D images (views) were formed for each angle of view using the rotation of the object.

Fig. 13. Original 2D (left) and depth (right) images.

The number of images is determined by the size of the screen point (pixel size) and the size of each element of the lenticular

$$N = \text{ceil}(LPI/\text{PixelSize}),$$

where ceil is the operation of rounding up, $LPI = 25.4/F$, F is the frequency of lenticular (number of elements per inch), PixelSize — the size of the screen point.

The round-up operation is used to obtain an integer value for the number of images. An autostereoscopic I_s image is created from an $M_s \times N_s \times 3$ matrix corresponding to the screen resolution and RGB color representation (Fig. 14).

Fig. 14. Schematic representation of 3D image file creation method.

The original image is converted into a matrix with a given resolution. All images, including I_s , are divided into segments of size $N \times N_s$. Each segment is a set of columns under one cylindrical lenticular lens. Each column with

number k is copied from the image corresponding to its angle of view. Since the sizes of the integer number of pixels and the lenticular LPI do not coincide, the following operation is applied: if less than half of the pixel size remains, the number of pixels under one lenticular is $N-1$, if more, then N . For this, N is rounded up, as noted above. The autostereoscopic image is composed of these segments, the number of which is M_s/N . Accordingly the resolution of the image for each angle of view is $M_s/N \times N_s$.

Fig. 15. 3D image files with 40LPI (left) and 70LPI (right) lenticular screens.

Fig. 16. 3D image files with 40LPI (left) and 70LPI (right) lenticular screens.

3.2. Experimental results

Below the images using mobile phones 4K Sony and LG with IPS LCD displays are presented. Lenticular sheets with 40 LPI and 70 LPI of different thicknesses were used in the experiments (Figs. 15 and 16).

It is followed from the experiments that the greater the viewing angle and view numbers, the less the image resolution and the depth range. The image resolution can be increased using the smartphones with higher resolution. To increase the depth of images the lenticular arrays with aspheric surface profiles should be developed.

An important step is a pitch test. The need for pitch test is to avoid miss-alignment between the lenses on the lenticular sheet and the interlaced image that is ready for displaying. We prepared pitch-test files for different lenticular screens (Fig. 17).

Fig. 17. Pitch-test file images for 40LPI (left) and 70LPI (right) lenticular screens.

An integral imaging method and capturing an object from different directions using a rotating table with a conventional camera can also be used to create 3D image files (Fig. 1). The drawback of the integral imaging method for capturing is the small depth. Capturing an object from different angles with the camera allows for a higher depth.

Below 3D images are presented when 3D files were created from 2D files obtained using a rotating table to capture images. A system consisting of a mobile phone with IPS LCD display matrix and a lenticular lens sheet with 70 LPI (lens pitch is 363 μm) has been developed. Note that a 3D display system using a mobile phone with a diamond pentile OLED display panel has been demonstrated in [23]. In [24] the 3D display consisting of a 2D OLED display panel, a pixel mask, and a lenticular lens was proposed.

In Figs. 18 and 19 the images from different viewing directions are shown. Here we used LG G3 mobile phone (resolution 2560×1440 pixels). 3D images with 8 views were created. 3D image has the ability to «look around». Objects (LED source, located behind the foot of the fighter), that are hidden for viewing in the image on the right, become visible from a different angle of view (left image).

Fig. 18. 3D images on the LG G3 mobile phone (2560×1440 pixels) screen from different viewing directions.
(a) The picture shows a red LED, which is located behind the foot of the warrior.

Fig. 19. 3D images on the LG G3 mobile phone (2560×1440 pixels) screen from different viewing directions.
Arrows show the relative positions of different parts of fighters.

In Fig. 19 the images using a LG G3 (2560×1440 pixels) and a 4K Sony Xperia XZ Premium (3840×2160 pixels) mobile phones and lenticular lens with 70 LPI are presented. 3D images with 8 and 12 views were created.

Although the resolution of 4K Sony mobile phone is higher than LG G3 phone, the picture quality is better when using LG phone. This is because the number of views under each lenticular lens for the LG phone ($N = \text{lens pitch} / \text{pixel size} = 7.9$) is closer to an integer value than for the 4K Sony phone ($N = 12.5$).

Fig. 20. 3D images on the LG G3 (a, b) and 4K Sony (c, d) mobile phones from different viewing directions.

Conclusions

Thus, optical design of display screens (lenticular array and diffusers) is performed and the algorithm for creation of 3D digital files for displaying images using multi-projectors and mobile phones is developed. The additional viewpoints, i.e. motion parallax give the necessary perception set of the 3D-scenes perspectives and have the property of “looking around” an object.

3D image files are created based on the depth and color data (2D+z files) and using capturing system based on the rotating table. It is shown, that the depth and view angle of 3D images are much better when using 3D files obtained with the help of a rotating table.

Matlab codes for 3D image file creation and combining images from four projectors are developed. Different screens consisting of lenticular sheets and diffuser layers are designed and experimental demonstrations of the 3D system are conducted.

Further developments we are considering in the creation of two-dimensional arrays with aspherical surface profiles of microlenses and increase in number of pixels. This will allow us significantly increase the view angle, depth range and resolution. The synthesis method including wave-optics and ray-tracing for the acceleration of the simulation of microlenses is effective to design microlenses with aspherical surfaces [25]. The holographic diffusers to increase the contrast and brightness of the images can be used in the display screen [26]. Experimental demonstration of 3D system, including image capturing, creation of 3D file and image displaying in real time will have significant importance for elucidation of possibilities of the technologies of digital processing and displaying of 3D images.

Proposed displays have a wide range of potential applications including 3D TVs and projection systems, mobile phones, as well as the systems for videoconference and medical applications.

Acknowledgements: The financial support of this research by the Russian Science Foundation (project No. 17-19-01461) is highly appreciated.

Список литературы

- Geng, J. Three-Dimensional Display Technologies // *Advances in Optics and Photonics*. 2013. Vol. 5. P. 456–535. DOI: 10.1364/AOP.5.000456
- Hong, J., Kim, Y., Choi, H.J., Hahn, J., et al. Three-Dimensional Display Technologies of Recent Interest: Principles, Status, and Issues // *Applied Optics*. 2011. Vol. 50. P. 87–115.
- Lueder, E. 3D Displays. UK: John Wiley & Sons, 2012. ISBN: 978-1-119-99151-9
- Jang, J. S., and Javidi, B. Real-Time All-Optical Three-Dimensional Integral Imaging Projector // *Appl. Opt.* 2002. Vol. 41. P. 4866–4869. DOI: 10.1364/AO.41.004866
- Liao, H., Iwahara, M., Hata, N., and Dohi, T. High-Quality Integral Videography Using a Multiprojector // *Opt. Express*. 2004. Vol. 12. P. 1067–1076. DOI: 10.1364/OPEX.12.001067
- Jang, J. S., and Javidi, B. Three-Dimensional Projection Integral Imaging Using Micro-Convex-Mirror Arrays // *Opt. Express*. 2004. Vol. 12(6). P. 1077–1083.
- Kim, Y., Park, S. G., Min, S. W., and Lee, B. Projection-Type Integral Imaging System Using Multiple Elemental Image Layers // *Appl. Opt.* 2011. Vol. 50. B18–B24. DOI: 10.1364/AO.50.000B18
- Jang, J. Y., Shin, D., Lee, B. G., and Kim, E. S. Multi-Projection Integral Imaging by Use of a Convex Mirror Array // *Opt. Lett.* 2014. Vol. 39. P. 2853–2856. DOI: 10.1364/OL.39.002853
- Takaki, Y., and Nago, N. Multi-Projection of Lenticular Displays to Construct a 256-View Super Multi-View Display // *Opt. Express*. 2010. Vol. 18(9). P. 8824–8835. DOI: 10.1364/OE.18.008824
- Takaki, Y., Takenaka, H., Morimoto, Y., Konuma, O., and Hirabayashi, K. Multi - View Display Module Employing MEMS Projector Array // *Opt. Express*. 2012. Vol. 20. P. 28257–28266. DOI: 10.1364/OE.20.028257
- Lee, J. H., Park, J., Nam, D., Choi, S. Y., Park, D. S., and Kim, C. Y. Optimal Projector Configuration Design for 300-Mpixel Multi-Projection 3D Display // *Opt. Express*. 2013. Vol. 21. P. 26820–26835. DOI: 10.1364/OE.21.026820
- Eldes, O., Aksit, K., and Urey, H. Multi-View Autostereoscopic Projection Display Using Rotatory Screen // *Opt. Express*. 2013. Vol. 21. P. 29043–29054. DOI: 10.1364/OE.21.029043
- Wang, Z., Wang, A., Wang, S., Ma, X., and Ming, H. Resolution-Enhanced Integral Imaging Using Two Micro-Lens Arrays with Different Focal Lengths for Capturing and Display // *Opt. Express*. 2015. Vol. 23. P. 28970–28977. DOI: 10.1364/OE.23.028970
- Moon, S., Park, S. G., Lee, C. K., Cho, J., Lee, S., and Lee, B. Computational Multi-Projection Display // *Opt. Express*. 2016. Vol. 24. P. 9025–9037. DOI: 10.1364/OE.24.009025
- Petrov, N. I., Sokolov, Y. M., Khromov, M. N., and Storozheva, A. L. Integral Imaging Multi-View 3D Display // *Frontiers in Optics/Laser Science Conference (FiO/LS)*. 2017. Washington, USA, Paper JT2A.107. DOI: 10.1364/FIO.2017.JT2A.107
- Petrov, N. I., Khromov, M. N., and Sokolov, Y. M. Large-Screen Multi-View 3D Display // *OSA Continuum*. 2019. Vol. 2(9). P. 2601–2613. DOI: 10.1364/OSAC.2.002601
- Okaichi, N., Miura, M., Sasaki, H., Watanabe, H., Arai, J., Kawakita, M., Mishina, T. Continuous Combination of Viewing Zones in Integral Three-Dimensional Display Using Multiple Projectors // *Opt. Eng.* 2018. Vol. 57(6). P. 061611. DOI: 10.1117/1.OE.57.6.061611
- Kim, J., et al. Crosstalk-Reduced Dual-Mode Mobile 3D Display // *J. Display Technol.* 2015. Vol. 11. P. 97–103.
- Markman, A., Wang, J., and Javidi, B. Three-Dimensional Integral Imaging Displays Using a Quick-Response Encoded Elemental Image Array // *Optica*. 2014. Vol. 1(5), 332–335. DOI: 10.1364/OPTICA.1.000332
- Woods, A. J. Crosstalk in Stereoscopic Displays: a Review // *J. Of Electronic Imaging*. 2012. Vol. 21. P. 040902. DOI: 10.1117/1.JEI.21.4.040902
- Woods, A. J., Harris, C. R., Leggo, D. B., and Rourke, T. M. Characterizing and Reducing Crosstalk in Printed Anaglyph Stereoscopic 3D Images // 2013. *Opt. Eng.* 2013. Vol. 52. P. 043203. DOI: 10.1117/1.OE.52.4.043203
- Son, J. Y., Lee, B. R., Park, M. C., and Lepotier, T. Crosstalk in Multiview 3-D Images // *Proc. SPIE*. 2015. Vol. 9495. Paper 94950P. DOI: 10.1117/12.2180110
- Algorri, J. F., Pozo, V. U., Sanchez-Pena, J. M., and Oton, J. M. An Autostereoscopic Device for Mobile Applications Based on a Liquid Crystal Microlens Array and an OLED Display // *J. Display Technol.* 2014. Vol. 10. P. 713–720. DOI: 10.1109/JDT.2014.2313143
- Ly, G. J., Zhao, B. C., Wu, F., and Wang, Q. H. Three-Dimensional Display with Optimized View Distribution // *Opt. Eng.* 2019. Vol. 58(2). P. 023108. DOI: 10.1117/1.OE.58.2.023108
- Petrov, N. I., and Petrova, G. N. Diffraction of Partially-Coherent Light Beams by Microlens Arrays // *Opt. Express*. 2017. Vol. 25(19), 22545–22564. DOI: 10.1364/OE.25.022545
- Petrov, N. I. Holographic Diffuser with Controlled Scattering Indicatrix // *Comp. Opt.* 2017. Vol. 41. P. 831–836. DOI: 10.18287/2412-6179-2017-41-6-831-836

Поступила 24 июня 2020 г.

ЗД ДИСПЛЕИ БЕЗ ОЧКОВ НА ОСНОВЕ ПРОЕКТОРОВ И МОБИЛЬНЫХ ТЕЛЕФОНОВ

Н.И. ПЕТРОВ, Ю.М. СОКОЛОВ, М.Н. ХРОМОВ

doi: 10.25210/jfop-2002-060073

Существуют различные типы 3D-систем отображения изображений, в том числе многовидовые, голографические и интегральные методы визуализации. 3D-дисплей, использующий интегральный метод, является одним из перспективных дисплеев, которые обеспечивают различные виды в зависимости от направления просмотра. Интегральная визуализация - это самый простой и подходящий метод создания реального 3D-изображения. В настоящей работе реализована и экспериментально продемонстрирована интегральная схема визуализации в режиме реального времени. Разработана многовидовая 3D-дисплейная система, включающая модули съемки трехмерных объектов, обработки изображений (создания файлов 3D-изображений) и оптический экран для воспроизводства 3D-изображений на основе интегральной технологии визуализации. Созданы большие экраны дисплея с использованием набора проекторов. Набор проекторов используется для улучшения характеристик изображения, таких как разрешение, глубина и угол обзора. Демонстрированы 3D изображения с использованием набора Full HD проекторов и проектора Ultra HD 4K. Показана возможность наблюдения скрытых объектов, расположенных за передним элементом, с использованием проекторов и смартфонов 4K Sony (разрешение 3840×2160), а также Samsung и LG (разрешение 2560×1440).

Практический интерес представляет перенос этих 3D-технологий на смартфоны и другие мобильные устройства. Демонстрированы 3D изображения на смартфонах 4K Sony (разрешение 3840×2160 пикселей), Samsung и LG (разрешение 2560×1440 пикселей). Для создания файлов 3D-изображений используются съемка объекта с разных сторон с помощью вращающейся платформы и обычной камеры. Недостатком интегрального метода для съемки является малая глубина. Съемка объекта с разных углов с помощью фотокамеры позволяет получить более высокую глубину изображения.

Предлагаемые подходы к воспроизводству 3D-изображений имеют широкий спектр потенциальных применений, включая 3D-телевизоры и проекционные системы, мобильные телефоны, а также системы видеоконференцсвязи и медицинские приложения.

Уважаемые читатели!

Подписка на журнал оформляется через подписное агентство «Книга-Сервис» и объединенный каталог «Пресса России» (подписной индекс **29196**)

Журнал включен:

- В *Российский индекс научного цитирования (РИНЦ)*.
- В базу данных *Russian Science Citation Index (RSCI)* на платформе *Web of Science*.
- В *Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук*.

МЕТОДЫ РАДИОЛОКАЦИОННЫХ И РАДИОМЕТРИЧЕСКИХ ИЗМЕРЕНИЙ

УДК 621.396: 537.87

ОСОБЕННОСТИ РЕШЕНИЙ ЗАДАЧ РАДИОЛОКАЦИИ ПРИ ИСПОЛЬЗОВАНИИ СВЕРХШИРОКОПОЛОСНЫХ СИГНАЛОВ

© Авторы, 2020

doi: 10.25210/jfor-2002-074081

Лаговский Б. А. — д.т.н., проф., МИРЭА — Российский технологический университет, Москва.

E-mail: robertlag@yandex.ru

Самохин А. Б. — д.ф.-м.н., проф., заслуженный деятель науки РФ, проф. кафедры прикладной математики, МИРЭА — Российский технологический университет. E-mail: absamokhin@yandex.ru

Аннотация

Представлены аналитические и численные решения задач повышения дальности действия сверхширокополосных РЛС за счет оптимизации их характеристик с учётом взаимовлияния отдельных излучателей. Обосновано использование новых эффективных алгоритмов обработки UWB-сигналов. Исследованы решения вариационных задач по поиску оптимальной формы зондирующих UWB-импульсов, обеспечивающих наибольшую вероятность обнаружения целей. Результаты аналитических и численных исследований показали, что оптимизация зондирующих UWB-импульсов позволяет в 2-4 раза повысить дальность действия наносекундных РЛС и одновременно улучшить угловые характеристики обнаружения целей.

Ключевые слова: сверхширокополосные сигналы, UWB радар, вариационные задачи радиолокации, дисперсионные характеристики антенных устройств

Abstract

Analytical and numerical solutions to problems of increasing the range of ultra-wide-band radars by optimizing their characteristics, taking into account the interaction of individual emitters, are presented. The use of new effective algorithms for processing UWB signals is justified. Solutions to variational problems of searching for the optimal form of probing UWB pulses that provide the highest probability of detecting targets are investigated. The results of analytical and numerical studies have shown that optimization of probing UWB pulses can increase the range of nanosecond radars by 2-4 times and simultaneously improve the angular characteristics of target detection.

Keywords: ultra-wideband signals, UWB radar, variational radar problems, dispersion characteristics of antenna devices

Введение

Интерес к применению UWB (ultra wide band), т.е. сверхширокополосных сигналов возник еще в 1960-х годах, когда появились первые работы по генерации и распространению ультракоротких импульсов, а также шумоподобных сигналов.

Во многих странах: в США, России, Китае, Франции, Японии и других активно ведут работы в области теории UWB технологий и создания на их основе образцов систем связи и радиолокации [1–5]. В настоящее время наиболее развитая область применения — обнаружение объектов на небольших расстояниях, составляющих единицы и десятки метров.

UWB-локация имеет ряд преимуществ по сравнению с обычной:

- значительно повышается точность измерений и разрешающая способность по дальности, достигающая десятков и даже единиц сантиметров;
- сигналы, поступающие от сложных целей представляют собой пачки импульсов отражённых от её отдельных элементов. Обработка поступающей информации позволяет получать радиоизображения этих целей и надёжно распознавать класс и тип исследуемых объектов,
- повышается вероятность обнаружения малозаметных объектов, защищённых радиопоглощающими покрытиями, в том числе выполненными по технологии «стелс», за счёт включения в спектр UWB-сигналов низкочастотных составляющих, которые слабо поглощаются покрытиями;
- повышается скрытность работы РЛС вследствие отсутствия боковых лепестков диаграммы направленности (ДН) при излучении.
- существенное уменьшение ЭПР (эффективной площади рассеяния) пассивных помех — дождя, тумана, аэрозолей и т.п., что облегчает наблюдение целей на их фоне,
- снижение размеров «мертвой зоны» РЛС, что позволяет создавать высокоточные радиодальномеры для систем посадки летательных и космических аппаратов, систем стыковки и т.д.
- обнаружение объектов в почве, в стенах, за листовым покровом, людей за препятствиями (в завалах, за стенами зданий), при медицинском диагностировании.

Для расширения области применения UWB-радаров необходимо значительно увеличить их дальность действия. Естественный путь решения — разработка более мощных генераторов. Это направление успешно развивается.

Другое направление повышения дальности действия — оптимизация характеристик UWB-радаров без увеличения энергетики радиоканала. Сверхширокая полоса используемых частот позволяет, как будет показано, значительно повысить эффективность решений задач оптимизации по сравнению с обычной радиолокацией и существенно увеличить дальность действия систем.

Рассмотрим ряд задач оптимизации характеристик передачи и приёма UWB-сигналов и создания эффективных алгоритмов обработки принимаемых сигналов.

Характеристики UWB-радаров во многом принципиально отличаются от характеристик РЛС других типов. Различия возникают из-за наличия дисперсии, пренебрежимо малой для узкополосных систем, но значительной в сверхширокой полосе используемых частот.

Выделим основные отличия:

- изменение формы и спектра сигнала. При использовании UWB-сигналов их временная структура, амплитудный и фазовый спектры существенно изменяются при излучении, при отражении сигнала от цели и при его приеме;
- зависимость ширины главного луча $\theta_{0,5}$, коэффициента направленного действия (КНД), уровня боковых лепестков антенной системы от временной структуры и длительности (или, при частотном анализе, от ширины спектра и распределения спектральной плотности) сигнала;
- зависимость формы (и спектра) излучаемого и принимаемого сигнала от углового расстояния до оси ДН;
- зависимость максимального значения, формы (и спектра) принимаемого UWB-сигнала от дисперсионных характеристик ЭПР цели;
- изменение основного уравнения радиолокации — уравнения дальности действия РЛС, т.к. ряд величин, входящих в уравнение, зависит от формы (и спектра) сигнала.

Аналитические исследования и результаты численных экспериментов на математической модели показывают, что изменения временной структуры (и спектра) UWB импульсов значительны, и ими нельзя пренебрегать при решении различного рода задач, связанных с излучением и приёмом сигналов.

Так, например, у UWB-импульса, излучаемого в направлении максимума ДН, по сравнению с генерируемым сигналом заметно увеличивается длительность, появляются дополнительные осцилляции, максимум спектра смещается в сторону высоких частот. Описанные изменения возникают из-за дисперсионных свойств ДН и КНД.

Для антенных решёток (АР), кроме того, временная структура излучаемых UWB-импульсов изменяется из-за дисперсионных свойств взаимовлияния элементов АР. Метод расчёта взаимных сопротивлений при излучении UWB-сигналов изложен в работе [5]. Если для высокочастотных составляющих UWB-импульса взаимовлияние пренебрежимо мало, то на низких частотах его необходимо учитывать при расчётах во избежание значительных ошибок.

Таким образом, дисперсию необходимо учитывать при расчётах характеристик и проведении оптимизации сверхширокополосных приёмо-передающих систем РЛС, а также при построении устройств обработки принимаемых сигналов.

1. Основные характеристики излучения и приёма UWB-сигналов

Дисперсионные зависимости основных характеристик работы UWB-РЛС от формы (спектра) излучаемого сигнала не позволяют ставить и решать задачи их оптимизации отдельно (например, отдельно для каждой из систем — формирования сигнала, излучения, приёма). Необходимо разработать новые аналитические и численные методы решения обратных и вариационных задач оптимизации наносекундных РЛС, учитывающие возникшие взаимосвязи основных характеристик. Частично такой анализ проведен в работах [2–4,7,8].

Определяя ДН по мощности $F(\theta, \gamma)$ как функцию, описывающую угловое распределение излучаемой мощности в пространстве по угловым координатам (θ, γ) приходим к выражению

$$F(\theta) = \int_0^\infty |V_e(\omega, \theta)|^2 d\omega = \int_0^\infty |V(\omega) f(\omega, \theta)|^2 d\omega, \quad (1)$$

где $V_e(\omega)$ — спектр излучаемого сигнала, $V(\omega)$ — спектр подаваемого на вход антенны UWB импульса, $f(\omega, \theta)$ — ДН излучающей системы на частоте ω .

Таким образом, при излучении сверхширокополосных сигналов форма ДН определяется не только свойствами антенны, но и шириной и формой спектра сигнала $V(\omega)$.

Ширина ДН по координате $\theta - \theta_{0,5}$, определяемая по уровню уменьшения излучаемой мощности в два раза, может быть найдена из решения уравнения

$$\int_0^\infty |V_e(\omega, \theta_0)|^2 d\omega - 0,5 \int_0^\infty |V_e(\omega, 0)|^2 d\omega = 0 \quad (2)$$

относительно θ_0 , и составляет $\theta_{0,5} = 2\theta_0$.

При отражении от цели комплексный спектр UWB-импульса может претерпеть значительные изменения, и для получения полной картины отражения UWB-сигнала описание цели с помощью ЭПР оказывается недостаточным. Необходимо учитывать также фазовые характеристики отражения.

Будем описывать отражающие свойства исследуемого объекта с помощью комплексной функции R , имеющей смысл коэффициента отражения. Функция R зависит от частоты зондирующего сигнала и углового расположения цели (α, γ) относительно РЛС

$$R(\alpha, \gamma, \omega) = \sqrt{\sigma_N(\alpha, \gamma, \omega)} \exp(i\phi(\alpha, \gamma, \omega)),$$

$$\sigma_N(\alpha, \gamma, \omega) = \frac{\sigma(\alpha, \gamma, \omega)}{\sigma_{\max}(0, 0, \omega)}, \quad (3)$$

где $\phi(\alpha, \gamma, \omega)$ — фазочастотная характеристика отражения, $\sigma_N(\alpha, \gamma, \omega)$ — нормированное на максимум по используемым частотам значение ЭПР — $\sigma(\alpha, \gamma, \omega)$.

В дальнейшем для краткости будем считать, что цель расположена в направлении $(0, 0)$ и обозначать введённые в (3) величины как $\sigma(\omega)$, $\phi(\omega)$, а также $R(\omega)$. Тогда спектр отражённого сигнала оказывается равным $V_R(\omega) = V_e(\omega)R(\omega) = V(\omega)f(\omega, 0, 0)R(\omega)$, а спектр принимаемого сигнала

$$V_r(\omega) = V_R(\omega)f(\omega, 0, 0) = V(\omega)f(\omega, 0, 0)^2 R(\omega). \quad (4)$$

Ширина ДН $\theta_{0,5}$ по углу θ , определяемая по уменьшению принимаемой мощности в два раза, находится из решения уравнения

$$\int_0^\infty |V_r(\omega, \theta_0)|^2 d\omega - 0,5 \int_0^\infty |V_r(\omega, 0)|^2 d\omega = 0. \quad (5)$$

Значение $\theta_{0,5} = 2\theta_0$, получаемое при решении (2) или (5), зависит не только от ДН, но и от спектра (или формы) UWB-сигнала. Решения (2) и (5) в общем случае может быть найдено только численно.

Используя введённые определения характеристик при использовании UWB-сигналов (1–5), перейдём к решению конкретных задач.

2. Поиск оптимальной формы UWB-импульсов

Определить форму генерируемого UWB-импульса $U(t)$ длительностью τ , со спектром, отличным от нуля в полосе частот $\omega \in [\omega_{\min}, \omega_{\max}]$, который обеспечивает наибольшую дальность обнаружения радиолокационной цели с заданной характеристикой отражения $R(\omega)$ при фиксированной полной энергии сигнала

$$\int_0^\tau U^2(t)dt = \int_{\omega_{\min}}^{\omega_{\max}} V(\omega)V^*(\omega)d\omega = \text{const}, \quad (6)$$

где V^* — комплексно-сопряжённая функция.

Заданной также считаем дисперсионную зависимость игольчатой, симметричной по углу отклонения от максимума ДН антенной системы $f(\omega, \theta)$.

Пусть функция $U_r(\theta, t)$ описывает зависимость от времени импульса, принимаемого с направления θ . Главным параметром при обнаружении UWB-импульса является его максимальное (пиковое) значение $U_M(\theta)$

$$U_M(\theta) = \max_t(U_r(\theta, t)) = \max_t \left[\int_{\omega_{\min}}^{\omega_{\max}} V(\omega)f_e(\omega, \theta)R(\omega)f_r(\omega, \theta)\exp(i\omega t)d\omega \right]. \quad (7)$$

Решая с помощью введения неопределённого множителя Лагранжа вариационную задачу (7) с дополнительным условием (6), получим с точностью до несущественной константы комплексный спектр оптимального UWB-импульса в виде

$$V_{opt}(\omega) = R^*(\omega) f^{*2}(\omega, \theta) \quad (8)$$

и соответствующую ему форму импульса

$$U_{opt}(\theta, t) = F^{-1}[R^*(\omega) f^{*2}(\omega, \theta)], \quad (9)$$

где $F^{-1}[\]$ — обратное преобразование Фурье.

Тогда принимаемый сигнал при излучении оптимального UWB-импульса (9) имеет вид

$$U_R(\theta, t) = F^{-1}\left[R(\omega)R^*(\omega)(f(\omega, 0)f^*(\omega, \theta))^2\right]. \quad (10)$$

Все спектральные составляющие принимаемого сигнала, как видно из (10), оказываются синфазными, что обеспечивает максимально возможный уровень отражённого полезного сигнала.

Работа системы в режиме использования оптимальных UWB-импульсов напоминает обработку сигнала с помощью оптимального фильтра (ОФ). Роль ОФ играет комплексный коэффициент отражения радиолокационной цели. Однако, в отличие от оптимальной фильтрации, шум при приёме сигнала не претерпевает изменений и остаётся белым. Это даёт возможность при дальнейшей обработке сигнала в приёмном устройстве увеличить отношение сигнал/шум за счёт применения ОФ.

Параметры принятого импульса сильно отличаются от излученного. Учитывая это обстоятельство, ОФ в приёмнике должен иметь частотную характеристику комплексно сопряжённую не со спектром излучаемого сигнала, как это осуществляется в узкополосных системах, а со спектром принимаемого импульса.

Такая обработка позволяет при приёме получить UWB-сигнал в виде

$$U_r(t, \theta) = F^{-1}\left[\left|R(\omega)R^*(\omega)\right|^4 \left(f(\omega, \theta)f^*(\omega, \theta)\right)^4\right] \quad (11)$$

и дополнительно увеличить значение максимума сигнала и отношение сигнал-шум.

Применение оптимальных UWB-импульсов и их согласованная фильтрация имеет ещё одно преимущество. Спектр излученного сигнала, а значит и его форма, зависят от углового расположения цели относительно РЛС — θ . Если форма импульса выбрана оптимальной для цели, расположенной в направлении максимума излучения $\theta = 0$, то для других направлений она оказывается неоптимальной. Тогда при приёме сигнала с направления θ , особенно после согласованной фильтрации, предназначенной для приёма с направления $\theta = 0$, максимальная величина импульса уменьшается с увеличением θ существенно быстрее, чем изменение значения ДН. Этот эффект эквивалентен уменьшению ширины ДН, что позволяет увеличить точность измерения угловых координат цели и угловое разрешение.

Для объектов, частично поглощающих падающие волны, в том числе для целей, защищённых радиопоглощающими покрытиями, применение оптимальных UWB-импульсов кроме значительного выигрыша в значении максимума $U_M(\theta)$, дополнительно на 20–30% увеличивает общую отражаемую энергию, что может быть использовано в задачах полуактивной радиолокации.

Примеры. В качестве примеров на рис. 1, 2 показаны изменения в спектрах и форме UWB-импульсов при излучении и приёме сигналов квадратной эквидистантной антенной решёткой с 41×41 излучателями и расстоянием $d = 0.7\lambda$ между элементами.

Рис. 1. Форма UWB-импульсов: 1 — генерируемый сигнал, 2 — излучаемый и 3 — принимаемый импульсы.

Объект располагался в направлении $\theta = 0$ и описывался комплексной характеристикой отражения $R(\omega)$ (3) в виде возрастающей с частотой функции с осцилляциями.

Представлены результаты, полученные при использовании биполярного UWB-импульса с амплитудным спектром, отличным от нуля в заданном частотном диапазоне $\omega \in [\omega_{\min}, \omega_{\max}]$.

Рис. 2. Амплитудные спектры UWB-импульсов: 1 — генерируемый, 2 —излучаемый и 3 — принимаемый.

Форма UWB-импульса претерпевает значительные изменения при излучении и приёме, а особенно при отражении от радиолокационной цели. Максимумы спектров смещаются в сторону возрастания частот, что объясняется повышением направленности антенных систем с ростом частоты. Изрезанность спектра возникает при отражении от сложных целей, в основном вследствие дисперсии фазовых характеристик отражения.

На рис. 3 представлена форма генерируемого оптимального UWB-импульса, полученная при решении Задачи 1. Для сравнения приведена форма биполярного импульса, использованного при получении результатов, показанных на рис. 1,2.

Энергия UWB-сигналов «размазана» по сверхширокому спектру, и спектральные составляющие сигналов этого типа оказываются сравнимы и даже ниже уровня шумов. Это приводит к значительным ошибкам при использовании частотных методик измерений. Как следствие, при расчётах, а особенно при измерениях, связанных с исследованием и использованием UWB-импульсов, удобно использовать временные характеристики сигналов и устройств.

В этом случае вместо ДН, ЭПР и всех других частотно зависимых характеристик РЛС и исследуемых объектов следует использовать их аналоги в виде импульсных характеристик. Тогда вместо коэффициента отражения $R(\omega)$ следует ввести импульсную характеристику отражения цели в виде Фурье-преобразования $h(t) = F[R]$.

Для иллюстрации на рис. 3 показана характеристика отражения $h(t)$ цели, использованной при решении задачи.

Импульсные характеристики взаимовлияния элементов АР, необходимые при решении задач оптимизации при использовании UWB-сигналов получены и описаны в [5].

Рис. 3. Временные зависимости. 1 — биполярный UWB-сигнал, 2 — импульсная характеристика отражения цели, 3 — оптимальный UWB-импульс.

Численные эксперименты показали, что при использовании оптимального UWB-импульса (9) значение максимума по мощности принимаемого сигнала увеличивается более, чем на два порядка по сравнению с импульсами типа 1 (см. рис. 3), сформированными без учёта дисперсионных характеристик цели.

На практике при формировании UWB-импульса, обеспечивающего наибольшую дальность действия системы, трудно рассчитывать на получение точной фазочастотной характеристики (ФЧХ) объекта. В этом случае целесообразно провести оценку ФЧХ основе какой-либо физической модели. Тогда, используя в виде $R(\omega)$ известную частотную зависимость ЭПР вместе с полученной модельной ФЧХ, можно провести частичную оптимизацию излучаемого сигнала, т.е. создать квазиоптимальный UWB-импульс.

Моделирование показало, что максимальные значения $U_m(\theta)$ при применении квазиоптимальных UWB-импульсов составляли $0.2 \div 0.5$ от значения оптимального (9).

Подводя итог, можно констатировать, что только за счёт оптимизации формы зондирующих UWB-импульсов, без увеличения энергетики канала, оказывается возможным увеличить дальность действия РЛС в 2–3 раза.

При обработке принимаемых сигналов широко используются ОФ и корреляторы. Их применение при цифровой обработке UWB-импульсов оказывается чрезвычайно эффективным из-за сверхширокой полосы сигнала. Так, для рассмотренного в примере сигнала использование ОФ с частотной характеристикой вида $V_r(\omega)^*$ (4) позволило повысить отношение сигнал/шум при приёме в 75 раз.

При UWB-локации ширина ДН в виде (2) и (5) зависит от ширины и формы спектра, а значит и от формы используемого UWB-сигнала. Следовательно, повышение дальности действия UWB-радара за счёт оптимизации формы импульса может привести к изменению углового разрешения и точности угловых измерений.

Для целей, расположенных в направлении $\theta \neq 0$, при приёме сигнала, оптимального для направления $\theta = 0$ вместо ожидаемого спектрального состава

$$V_0(\omega) = R(\omega)R^*(\omega)(f(\omega, 0)f^*(\omega, 0))^2 \quad (12)$$

получим

$$V_{np}(\omega, \theta) = R(\omega)R^*(\omega)(f(\omega, \theta)f^*(\omega, \theta))(f(\omega, 0)f^*(\omega, 0)) \quad (13)$$

и соответствующие (12,13) различные временные зависимости UWB-сигналов в виде $U_R(\theta, t) = F^{-1}[V]$.

Как уже указывалось, UWB-импульсы (9), (10) неоптимальны для направлений отличных от $\theta = 0$, и при их использовании максимальная величина принимаемых импульсов уменьшается с увеличением θ быстрее, чем значения ДН. Как следствие — эффективная ширина ДН, определяемая по максимальным значениям сигнала для каждого направления, уменьшается, что даёт возможность увеличить точность угловых измерений.

На рис. 4 приведены полученные на основе (12,13) нормированные эффективные ДН при приёме UWB-импульсов. Показана область углов, ограниченная главным лепестком.

Рис. 4. Эффективная диаграмма направленности при приёме: 1 — биполярного сигнала, 2 — оптимального UWB-импульса, 3 — оптимального UWB-импульса после прохождения ОФ.

Использование при приёме ОФ с частотной характеристикой $S(\omega) = V_0(\omega)$ (12), т.е. оптимального только для направления расположения цели $\theta = 0$, резко увеличивает максимальное значение UWB-импульса, принимаемого с этого направления. В итоге ширина эффективной приёмной ДН заметно снижается.

Из приведенных на рис. 4 результатов следует, что использование оптимальных для исследуемых целей UWB-импульсов позволяет не только повысить дальность действия UWB-радара, но и увеличить точность угловых измерений и угловое разрешение в 2–4 раза.

В рассмотренной задаче был получен максимально возможный сигнал, отраженный от цели с известными характеристиками отражения. Получаемая ширина ДН при этом не контролировалась. Для узкополосных систем она не меняется при изменении формы сигнала.

Эффект значительного увеличения ширины ДН, в частности, возникает, если цель на низких частотах характеризуется заметной ЭПР, а с увеличением частоты значение ЭПР резко снижается, вплоть до значений близких к нулю. Такая зависимость характерна для ЭПР целей, защищенных сверхширокополосными радиопоглощающими покрытиями.

Возникает задача поиска оптимального спектрального состава импульса, при использовании которого дальность действия РЛС оказывается максимально возможной и, одновременно, ширина ДН оказывается не больше заданной величины, т.е. угловые координаты определяются с необходимой точностью.

Решение этой и близких вариационных задач оптимизации в области UWB-радиолокации изложено в публикациях [6–9]. Продемонстрирована эффективность применения полученных решений.

Исследование выполнено при финансовой поддержке РФФИ в рамках проекта № 20–07–00006.

Заключение

Показано, что спектр и временная структура излучаемых и принимаемых антенными системами UWB-импульсов, значительно отличаются от спектра и временной структуры импульсов, подаваемых на излучатели из-за дисперсионных свойств ДН и взаимовлияния элементов АР, также носящего дисперсионный характер. Определена форма импульсов, обеспечивающих наибольшую дальность обнаружения цели с известными характеристиками отражения при заданной энергии излучаемого сигнала и фиксированной ширине спектра UWB-импульса. Результаты численных исследований показали, что оптимизация позволяет в 2–4 раза повысить дальность действия UWB-радара. Обоснована возможность оптимизации спектрального состава UWB-импульсов для решения задачи одновременного повышения дальности действия систем и улучшения угловых характеристик обнаружения целей.

Список литературы

1. Terrence, W. Barrett. History of UWB Radar and Communications: Pioneers and Innovators // Progress in Electromagnetics Symposium (PIERS). 2000. Microwave Journ. January 2001.
2. Holami, G., Mehrpourbernety, H., and Zakeri, B. UWB Phased Array Antennas for High Resolution Radars // Proc. Of the 2013 International Symposium on Electromagnetic Theory. 2013. P. 532–535. ISBN: 978–4–88552–277–2
3. Zhou Yuan, Law Choi Look, Xia Jingjing. Ultra Low-Power UWB-RFID System for Precise Location-Aware Applications // 2012 IEEE Wireless Communications and Networking Conference. Workshops (WCNCW). P. 154–158. DOI: 10.1109/WCNCW.2012.6215480
4. Khan, H.A., Edwards, D.J., and Malik, W.Q. Ultra Wideband MIMO Radar // Proc. IEEE Intl. Radar Conf. Arlington, VA, USA. 9 May 2005.
5. Lagovsky, B.A., Samokhin, A.B., Shestopalov, Y.V. // Pulse Characteristics of Antenna Array Radiating UWB Signals. Proc. Of the 10th European Conf. On Antennas and Propagation (EuCAP). 2016. Davos, Switzerland. P. 2479–2482. DOI: 10.1109/EuCAP.2016.7481624
6. Lagovsky, B., Samokhin, A., Samokhina, A. Image Restoration of the Multiple Target by Smart Antenna Array Radiating UWB Signals // Proceedings of Electromagnetics in Advanced Applications (ICEAA). 2015. Torino, Italy. P. 363–365. DOI: 10.1109/ICEAA.2015.7297136
7. Lagovsky, B.A., Chikina, A.G. Shape Optimization of UWB Pulses // Progress in Electromagnetics Research Symposium (PIERS). 2017. P. 475–478. DOI: 10.1051/itmconf/20193005017
8. Lagovsky, B.A., Samokhin, A.B., and Shestopalov, Y.V. Increasing the Range of UWB Radars // Proc. 2018 12th European Conf. On Antennas and Propagation (EUCAP). 2018. UK, Ser. IET Conference Publications. Vol. CP7 41. DOI: 10.1049/cp.2018.1255
9. Lagovsky, B.A., Samokhin, A.B., and Shestopalov, Y.V. Increasing Accuracy of Angular Measurements Using UWB Signals // 11th European Conf. On Antennas and Propagation (EUCAP). 2017. Paris. P. 1083–1086. DOI: 10.23919/EuCAP.2017.7928204

Поступила 10 июня 2020 г.

FEATURES OF SOLVING RADAR PROBLEMS USING ULTRA-WIDE-BAND SIGNALS

B.A. LAGOVSKY, A.B. SAMOKHIN
Russian Technological University (MIREA)

doi: 10.25210/jfop-2002-074081

Expanding the scope of UWB radars is possible with a significant increase in their range. Optimizing the characteristics of UWB radars without increasing the power of the radio channel allows us to partially solve this problem.

The ultra-wide band of frequencies used can significantly improve the efficiency of optimization solutions compared to conventional radar and, as a result, significantly increase the range of systems.

The problems of optimizing the form of UWB signals and creating effective algorithms for their processing are considered. It is shown that the spectrum and time structure of UWB pulses emitted and received by antenna systems differ significantly from the spectrum and time structure of pulses fed to emitters due to the dispersion properties of the antenna pattern and mutual impedances of the antenna array elements.

It is confirmed that the dispersion dependence of the main performance characteristics of UWB radars on the complex spectrum of the generated signal does not allow us to set and solve problems separately for each of the systems.

A new method for solving variational optimization problems for nanosecond radars is presented, taking into account the relationships between the main characteristics. The form of pulses that provide the greatest detection range of the target with known reflection characteristics at a given energy of the emitted signal and fixed width of the spectrum of UWB pulses is determined. The optimal filter characteristics for goals with the specified reflection characteristics are found. The efficiency of using pulse characteristics to describe the processes of radiation, reflection and reception of UWB signals is proved.

The results of analytical and numerical studies have shown that optimization of probing UWB pulses allows 2-4 times to increase the range of nanosecond radars.

The possibility of optimizing the spectral composition of UWB pulses to solve the problem of simultaneously increasing the range of systems and improving the angular characteristics of target detection is proved.

Уважаемые читатели!

Подписка на журнал оформляется через подписное
агентство «Книга-Сервис» и объединенный каталог
«Пресса России» (подписной индекс **29196**)

Журнал включен:

- В *Российский индекс научного цитирования (РИНЦ)*.
- В базу данных *Russian Science Citation Index (RSCI)* на платформе *Web of Science*.
- В *Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук*.

ЭЛЕКТРОДИНАМИКА ИСКУССТВЕННЫХ СРЕД И СТРУКТУР

УДК 530.182

ОБ УСЛОВИЯХ ВЗАИМОДЕЙСТВИЯ ВОЛН

© Авторы, 2020

doi: 10.25210/jfor-2002-082090

Буц В. А. — д.ф.м.н., проф., начальник лаборатории, Национальный научный центр «Харьковский физико-технический институт»; Институт радиоастрономии НАН Украины; Харьковский национальный Университет им. В.Н. Каразина, Харьков Украина. E mail: vbuts@kipt.kharkov.ua

Тарасов Д. В. — м.н.с., Национальный Научный центр «Харьковский физико-технический институт».

Аннотация

Описаны новые условия взаимодействия волн. Эти условия являются более общими, чем известные. Они, как частный случай, содержат известные, но дополнительно учитывают тот факт, что в распределенных системах расстройка вдоль одной из пространственно-временных осей может быть скомпенсирована расстройкой вдоль других осей. Рассмотрены примеры взаимодействия волн, характеристики которых удовлетворяют новым условиям. Рассмотренные примеры позволяют обнаружить новые условия в эксперименте.

Ключевые слова: взаимодействие волн в средах, трехволновое взаимодействие, нелинейная динамика, распадные процессы, обмен энергией между волнами

Abstract

The new conditions of wave interaction are described. These conditions are more general than the known conditions. They, as a special case, contain well-known conditions. New conditions reflect the fact that in distributed systems, the detuning along one of the spatiotemporal axes can be compensated by the detuning along other axes. Examples of the interaction of waves whose characteristics satisfy new conditions are considered. The considered examples allow to detect new conditions in the experiment.

Keywords: wave interaction in medium, three-wave interaction, nonlinear dynamics, decay processes, energy exchange between waves

Введение

Известно, что эффективное (резонансное) взаимодействие волн происходит, при выполнении условий:

$$\Delta \vec{k} = \sum_i \vec{k}_i = 0, \Delta \omega = \sum_i \alpha_i \omega_i = 0; \alpha_i = \pm 1. \quad (1)$$

Эти условия аналогичны условиям резонансного взаимодействия частиц. Для волн они означают, что должны выполняться условия синхронизма между взаимодействующими волнами вдоль каждой из четырех осей четырехмерного пространства.

Рис. 1. Процессы распада $\Delta \omega = \omega_k - \omega_{k_1} - \omega_{k_2} = 0$; $\Delta \vec{k} = \vec{k} - \vec{k}_1 - \vec{k}_2 = 0$ (а)

и слияния $\Delta \omega = \omega_{k_1} + \omega_{k_2} - \omega_k = 0$; $\Delta \vec{k} = \vec{k}_1 + \vec{k}_2 - \vec{k} = 0$ (б) при трехволновом взаимодействии.

Эти условия можно назвать идеальными условиями взаимодействия волн. Если появляется хотя бы одна отличная от нуля расстройка, то эффективность взаимодействия волн уменьшается. Это известные факты (см., например, [1–3]). Однако взаимодействие волн происходит не локально. Этот процесс распределен в некоторой пространственно-временной области. Поэтому можно ожидать, что имеются некоторые направления в этом пространстве (характеристические, субхарактеристические), в которых расстройка вдоль, например временной оси ($\Delta \omega \neq 0$), может быть скомпенсирована расстройкой вдоль одной из пространственных осей. В настоящей работе показано, что действительно такие направления существуют. Получены условия для эффективного взаимодействия волн при наличии расстроек.

В работе будут рассмотрены только трехволновые взаимодействия. Для них известные условия взаимодействия представлены на рис. 1 [1].

Ниже, в разделе 2, сформулирована задача. В третьем разделе рассмотрена наиболее простая задача о взаимодействии волн в среде с пространственно-временной периодической неоднородностью. Причем, пространственная неоднородность является одномерной. В разделе 4 эта линейная задача обобщена на случай трехмерной пространственной неоднородности. В пятом разделе рассмотрена нелинейная динамика трехволнового взаимодействия. Анализ простой модели, которая может быть реализована в эксперименте, содержится в шестом разделе. В заключении формулируются и обсуждаются основные результаты.

1. Постановка задачи. Основные уравнения

Рассмотрим наиболее простую модель взаимодействия волн, в которой могут реализоваться новые условия взаимодействия волн. Эта модель трехволнового взаимодействия, в которой одна из взаимодействующих волн является фиксированной волной. Характеристики этой волны не меняются. В качестве такой волны можно рассматривать диэлектрическую проницаемость среды. Например, рассмотрим среду, диэлектрическую проницаемость которой можно представить в виде

$$\varepsilon = \varepsilon_0 + \tilde{\varepsilon}, \quad \tilde{\varepsilon} = q \cos(\vec{k}\vec{r} - \Omega t), \quad q \ll 1. \quad (2)$$

Пусть в такой среде распространяется две электромагнитные волны, частоты которых различны. Будем интересоваться условиями эффективного взаимодействия этих волн в такой среде. Уравнения для каждой из этих волн — это уравнение Максвелла. Из этих уравнений легко найти уравнение для векторов электрического поля волн в такой среде

$$\Delta \vec{E} - \frac{1}{c^2} \frac{\partial^2 (\varepsilon \vec{E})}{\partial t^2} = -\vec{\nabla} \left(\frac{1}{\varepsilon} \vec{E} \cdot \vec{\nabla} \varepsilon \right). \quad (3)$$

По предположению, имеется две волны, поэтому решение уравнения (3) будем искать в виде суммы двух слагаемых

$$\vec{E} = \vec{A}_0(\vec{r}, t) \exp(-i\omega_0 t + i\vec{k}_0 \vec{r}) + \vec{A}_1(\vec{r}, t) \exp(-i\omega_1 t + i\vec{k}_1 \vec{r}). \quad (4)$$

Здесь

$$k_0^2 = \omega_0^2 \varepsilon_0 / c^2, \quad k_1^2 = \omega_1^2 \varepsilon_0 / c^2.$$

Каждое из слагаемых в (4) описывает плоскую электромагнитную волну, амплитуды которых являются функциями времени и координат. Так как пространственно-временная неоднородность среды предполагается малой, то эти амплитуды мы будем предполагать медленноменяющимися функциями пространства и времени.

2. Среда с одномерной неоднородностью

Рассмотрим наиболее простой случай, который проявляет наиболее важные характеристики новых условий взаимодействия и который, по-видимому, наиболее легко реализовать в эксперименте. Будем считать, что взаимодействие происходит между поперечными волнами, что среда периодически-неоднородна только в одном направлении (в направлении z). В этом случае временная расстройка может быть скомпенсирована только расстройкой вдоль оси z . Более того, будем предполагать, что волны расположены на одной дисперсионной линейной ветке. В этом случае фазовые скорости волн, и их групповые скорости совпадают. Как будет видно ниже, все эти ограничения являются непринципиальными (они наложены только для упрощения формул) и при необходимости могут быть убраны. Будем также считать, что волны распространяются только в одном направлении — в направлении оси z . В этом случае, подставляя (4) в (3), можно получить следующие уравнения для нахождения амплитуд $\vec{A}_i$:

$$\begin{aligned} \left[\frac{\partial^2 \vec{A}_0}{\partial z^2} - \frac{\varepsilon_0}{c^2} \frac{\partial^2 \vec{A}_0}{\partial t^2} \right] + 2i \left[k_0 \frac{\partial \vec{A}_0}{\partial z} + \frac{\varepsilon_0 \cdot \omega_0}{c^2} \frac{\partial \vec{A}_0}{\partial t} \right] &= -\frac{q (\omega_1 \pm \Omega)^2}{2 c^2} \vec{A}_1 \cdot \exp[-i \cdot \delta(z, t)] \\ \left[\frac{\partial^2 \vec{A}_1}{\partial z^2} - \frac{\varepsilon_0}{c^2} \frac{\partial^2 \vec{A}_1}{\partial t^2} \right] + 2i \left[k_1 \frac{\partial \vec{A}_1}{\partial z} + \frac{\varepsilon_0 \cdot \omega_1}{c^2} \frac{\partial \vec{A}_1}{\partial t} \right] &= -\frac{q (\omega_0 \pm \Omega)^2}{2 c^2} \vec{A}_0 \cdot \exp[i \cdot \delta(z, t)] \end{aligned} \quad (5)$$

где $\delta(\vec{r}, t) \equiv \Delta \vec{k} \cdot \vec{r} - \Delta \omega \cdot t$, $\Delta \vec{k} \equiv (\vec{k}_1 - \vec{k}_0 \pm \vec{k})$, $\Delta \omega \equiv \omega_1 - \omega_0 \pm \Omega$.

Удобно ввести следующие новые независимые переменные: $\xi = z - \tau$ и $\eta = z + \tau$, где $\tau = t \cdot c / \sqrt{\varepsilon_0}$.

Система (5) при этом переписывается в таком виде

$$\begin{aligned}\frac{\partial^2 A_0}{\partial \xi \cdot \partial \eta} + 2ik_0 \frac{\partial A_0}{\partial \eta} &= -\frac{q}{2} \frac{\omega_1^2}{c^2} A_1 \cdot \exp(i \cdot \delta) \\ \frac{\partial^2 A_1}{\partial \xi \cdot \partial \eta} + 2ik_1 \frac{\partial A_1}{\partial \eta} &= -\frac{q}{2} \frac{\omega_0^2}{c^2} A_0 \cdot \exp(-i \cdot \delta).\end{aligned}\quad (6)$$

Эта система строгая. Пусть расстройки частот и волновых векторов таковы, что характеристики уравнений (5) без вторых производных (т.е. субхарактеристики) параллельны прямым

$$\delta(\vec{r}, t) \equiv \Delta k \cdot z - \Delta \omega \cdot t = \text{const}(\eta).$$

Это означает, что производные вдоль этих субхарактеристик равны нулю $\partial \delta / \partial \eta = 0$.

Взаимодействие волн обусловлено малой неоднородностью ($q \ll 1$) диэлектрической проницаемости. Естественно ожидать, что амплитуды волн будут меняться медленно. Поэтому в системе уравнений (6) можно опустить вторые производные. Следует заметить, что это предположение всегда требует дополнительного анализа. В частности, как минимум, полученные решения следует проверять на удовлетворение этому предположению.

Отбросим вторые производные в левой части системы (6). Тогда левая часть этих уравнений может рассматриваться как производная вдоль характеристических прямых: $\eta = z + \tau = C = \text{const}$. Причем, эти направления для взаимодействующих волн совпадают. Кроме того, выполняются такие соотношения

$$\frac{\partial \delta}{\partial \eta} = 0 \quad \text{или} \quad \Delta \omega = \Delta k \cdot c / \sqrt{\varepsilon_0}. \quad (7)$$

Обратим также внимание на то, что правые части системы уравнений (6) содержат комплексно-сопряженные множители $\exp[i \cdot \delta(z, t)]$. Учитывая этот факт и соотношение (7) из системы уравнений (6), получим следующие уравнения для определения амплитуд A_0 и A_1

$$\frac{\partial^2 A_i}{\partial \eta^2} + K^2 \cdot A_i = 0, \quad i = \{0, 1\}, \quad (8)$$

где $K^2 = q^2 (k_0 \cdot k_1) / 64$.

Решениями уравнений (8) могут быть, например, функции $A_0 = a \cos(K\eta)$, $A_1 = b \sin(K\eta)$.

Такой выбор решений, в частности, может означать, что происходит периодическая перекачка энергии из одной волны во вторую волну в процессе их взаимодействия вдоль «направления» $\eta = z + \tau$.

Рассмотренное выше взаимодействие волн происходит в том случае, когда они распространяются в одном и том же направлении, в направлении положительной оси z ($k_0 > 0, k_1 > 0$). Легко показать, что если они будут распространяться в противоположном направлении ($k_0 < 0, k_1 < 0$), то взаимодействие будет происходить по тем же законам взаимодействия. Однако это взаимодействие будет происходить вдоль другой субхарактеристики. А именно вдоль $\xi = z - \tau$. В этом случае уравнение (8) будет преобразовано к уравнению

$$\frac{\partial^2 A_i}{\partial \xi^2} + K^2 \cdot A_i = 0, \quad i = \{0, 1\}. \quad (9)$$

Решениями этих уравнений могут быть функции $A_0 = a \cos(K\xi)$ и $A_1 = b \sin(K\xi)$.

3. Среда с трехмерной неоднородностью

Полученные выше результаты обобщаются на случай, когда среда имеет трехмерную неоднородность. Взаимное расположение волновых векторов взаимодействующих волн и их единичных векторов поляризации представлены на рис. 2. При этом, подставляя поля (4) в уравнение (3) после громоздких, но простых вычислений, можно получить следующую систему дифференциальных уравнений в частных производных первого порядка для определения динамики во времени и пространстве амплитуд $\vec{A}_0$ и $\vec{A}_1$

$$\begin{aligned}\frac{\partial \vec{A}_0}{\partial l_0} &= \frac{q \cdot i}{4N_0} \left\{ \frac{(\omega_1^2 + \Omega^2)}{c^2} \cdot \vec{A}_1 \pm \frac{1}{\varepsilon_0} (\vec{k}_1 \pm \vec{\kappa}) (\vec{\kappa} \vec{A}_1) \right\} \exp(i\delta), \\ \frac{\partial \vec{A}_1}{\partial l_1} &= \frac{q \cdot i}{4N_1} \left\{ \frac{(\omega_0^2 + \Omega^2)}{c^2} \cdot \vec{A}_0 \mp \frac{1}{\varepsilon_0} (\vec{k}_0 \mp \vec{\kappa}) (\vec{\kappa} \vec{A}_0) \right\} \exp(-i\delta).\end{aligned}\quad (10)$$

Здесь $\delta(\vec{r}, t) \equiv \Delta \vec{k} \cdot \vec{r} - \Delta \omega \cdot t$; $\Delta \vec{k} \equiv (\vec{k}_1 - \vec{k}_0 \pm \vec{\kappa})$; $\Delta \omega \equiv \omega_1 - \omega_0 \pm \Omega$; $N_i = \sqrt{k_{ix}^2 + k_{iy}^2 + k_{iz}^2 + \varepsilon \cdot \omega_i^2 / c^2}$ — норма

для каждой из волн. Производные в системе (10) взяты вдоль характеристических направлений: $\vec{l}_{0,1} = \{k_{0,1,x}; k_{0,1,y}; k_{0,1,z}; \omega_0 \cdot \sqrt{\varepsilon_0} / c\} / N_{0,1}$, где $\vec{l}_0$ и $\vec{l}_1$ — единичные вектора вдоль соответствующих характеристических направлений.

Уравнение $\delta(\vec{r}, t) = C = \text{const}(l_0, l_1)$ представляет собой уравнение гиперплоскости в четырехмерном пространстве $(\vec{r}, t)$. Чтобы эта расстройка не менялась в процессе обмена энергией между волнами, необходимо, чтобы характеристические прямые

$$\begin{aligned} x &= \alpha k_{0x} + x_0; y = \alpha k_{0y} + y_0; z = \alpha k_{0z} + z_0; t = \alpha \cdot (\omega_0 \cdot \varepsilon_0 / c^2) + t_0; \\ x &= \alpha_1 k_{1x} + x_1; y = \alpha_1 k_{1y} + y_1; z = \alpha_1 k_{1z} + z_1; t = \alpha_1 \cdot (\omega_1 \cdot \varepsilon_0 / c^2) + t_1 \end{aligned} \quad (11)$$

были параллельны гиперплоскости $\delta(\vec{r}, t) = C = \text{const}$. Этим условиям удовлетворяют соотношения

$$(\Delta \vec{k} \cdot \vec{k}_0) c^2 - \Delta \omega \cdot \omega_0 \cdot \varepsilon_0 = 0, \quad (\Delta \vec{k} \cdot \vec{k}_1) c^2 - \Delta \omega \cdot \omega_1 \cdot \varepsilon_0 = 0. \quad (12)$$

Взаимное расположение векторов $\Delta \vec{k}$, $\vec{k}_0$ и $\vec{k}_1$, которые удовлетворяют этим условиям, представлено на рис. 3. Легко проверить, что при выполнении условий (12), производные вдоль характеристических направлений $\vec{l}_0$ и $\vec{l}_1$ от расстройки будут равны нулю $\partial \delta(\vec{r}, t) / \partial l_0 = \partial \delta(\vec{r}, t) / \partial l_1 = 0$.

Рис. 2. Взаимное расположение векторов $\vec{k}_0$ и $\vec{k}_1$, а также единичных векторов поляризации $\vec{e}_i$.

Рис. 3. Взаимного расположения векторов $\Delta \vec{k}$, $\vec{k}_0$ и $\vec{k}_1$, которое необходимо для эффективного взаимодействия волн.

Связь амплитуд $\vec{A}_0$ и $\vec{A}_1$ в системе уравнений (10) более сложная, чем в системе (5). В средах с такими характеристиками различные поляризации оказываются, в общем случае, связанными. Поэтому амплитуды волн представим в виде суммы "σ" и "π" поляризации

$$\vec{A}_0 = \vec{e}_4 \cdot A_{01} + \vec{e}_2 \cdot A_{02}; \quad \vec{A}_1 = \vec{e}_5 \cdot A_1 + \vec{e}_3 \cdot A_2. \quad (13)$$

Здесь $\vec{e}_i$ — единичные вектора, определяющие направления поляризации электрических векторов. Направления этих векторов относительно волнового вектора $\vec{k}_0$ и волнового вектора $\vec{k}_1$ представлены на рис. 2. Вектора $\vec{e}_5$ и $\vec{e}_4$ параллельны и направлены перпендикулярно плоскости рисунка. Из этого рисунка, в частности, видно, что рассматривается взаимодействие (через неоднородную среду) двух

поперечных электромагнитных волн ($\vec{e}_2 \perp \vec{k}_0$, $\vec{e}_4 \perp \vec{k}_0$; $\vec{e}_3 \perp \vec{k}_1$, $\vec{e}_5 \perp \vec{k}_1$). Если выполняются обычные условия синхронизма $\Delta\vec{k} = 0$ и $\Delta\omega = 0$, то, вектор $\vec{k}$ лежит в одной плоскости с волновыми векторами $\vec{k}_0$ и $\vec{k}_1$. В этом случае, домножив первое уравнение системы (10) на $\vec{e}_4$, а второе уравнение системы (10) на $\vec{e}_5$ ($\vec{e}_4 \parallel \vec{e}_5 \perp \vec{k}_i$), получим следующую систему замкнутых уравнений первого порядка для определения амплитуд волн A_{01} и A_1

$$\begin{aligned}\frac{\partial A_{01}}{\partial l_0} &= \frac{q}{4 \cdot i \cdot N_0} \left\{ -\frac{\omega_0^2}{c^2} \cdot A_1 \right\}, \\ \frac{\partial A_1}{\partial l_1} &= \frac{q}{4 \cdot i \cdot N_1} \left\{ -\frac{\omega_1^2}{c^2} \cdot A_{01} \right\}.\end{aligned}\quad (14)$$

Система уравнений (14) описывает динамику "σ" - поляризации. Аналогично, домножив первое уравнение системы (10) на $\vec{e}_2$, а второе уравнение этой системы на $\vec{e}_3$, получим систему уравнений, описывающую динамику амплитуд A_{02} и A_2 ("π" — поляризация). Динамика взаимодействия волн в этом случае достаточно хорошо изучена.

В общем случае, когда $\Delta\vec{k} \neq 0$, получить замкнутую систему, описывающую динамику отдельно волн с "σ" поляризацией и волн с "π" -поляризацией, не удастся. Однако для ответа на интересующий вопрос об эффективном взаимодействии волн достаточно рассмотреть динамику не каждой компоненты электрического поля взаимодействующих волн, а достаточно рассмотреть динамику функций $(\vec{k} \cdot \vec{A}_0)$ и $(\vec{k} \cdot \vec{A}_1)$.

Для таких функций из системы (10) получим следующую систему уравнений:

$$\begin{aligned}\frac{\partial(\vec{k} \cdot \vec{A}_0)}{\partial l_0} &= \frac{q}{4 \cdot i \cdot N_0} \left\{ -\frac{(\omega_1 \pm \Omega)^2}{c^2} \pm \frac{\vec{k}}{\varepsilon_0} (\vec{k}_1 \pm \vec{k}) \right\} (\vec{k} \cdot \vec{A}_1) \cdot \exp(i\delta), \\ \frac{\partial(\vec{k} \cdot \vec{A}_1)}{\partial l_1} &= \frac{q}{4 \cdot i \cdot N_1} \left\{ -\frac{(\omega_0 \mp \Omega)^2}{c^2} \mp \frac{\vec{k}}{\varepsilon_0} (\vec{k}_0 \pm \vec{k}) \right\} (\vec{k} \cdot \vec{A}_0) \cdot \exp(-i\delta).\end{aligned}\quad (15)$$

В этом случае динамику функций $(\vec{k} \cdot \vec{A}_0)$ и $(\vec{k} \cdot \vec{A}_1)$ можно изучать независимо

$$\begin{aligned}\frac{\partial^2(\vec{k} \cdot \vec{A}_0)}{\partial l_0 \partial l_1} &= -\frac{q^2}{16 \cdot N_0 \cdot N_1} \left\{ -\frac{(\omega_1 \pm \Omega)^2}{c^2} \pm \frac{\vec{k}}{\varepsilon_0} (\vec{k}_1 \pm \vec{k}) \right\} \times \\ &\times \left\{ -\frac{(\omega_0 \mp \Omega)^2}{c^2} \mp \frac{\vec{k}}{\varepsilon_0} (\vec{k}_0 \pm \vec{k}) \right\} (\vec{k} \cdot \vec{A}_0).\end{aligned}\quad (16)$$

Аналогичное уравнение получается и для $(\vec{k} \cdot \vec{A}_1)$. Уравнения для функций $(\vec{k} \cdot \vec{A}_0)$ и $(\vec{k} \cdot \vec{A}_1)$ представляют собой линейные уравнения с частными производными второго порядка с постоянными коэффициентами.

Возникает вопрос: «Могут ли дать условия (12) что-либо нового по сравнению с известными условиями?» Ниже покажем, что действительно они дают новые возможности для взаимодействия волн.

Если среда стационарна, то $\Omega = \Delta\omega = 0$. Соответственно, процесс распространения волн в такой среде не будет зависеть от времени. При этом должно выполняться одно из привычных условий синхронизма — $\Delta\omega = 0$ (закон сохранения энергии). В этом случае из (12) следует, что кроме привычных условий синхронизма $\Delta\vec{k} = 0$, имеются дополнительные возможности для эффективного взаимодействия волн. Они заключаются в том, что волновой вектор $\vec{k}_0$ и волновой вектор $\vec{k}_1$ должны быть перпендикулярны вектору $\Delta\vec{k}$ (см. рис. 4). Анализ этих условий показывает, что возможны, например, следующие условия для эффективного взаимодействия волн. Первое заключается в том, что $\vec{k}_0 = \vec{k}_1$ и, по-видимому, является не очень интересным. Второе требует, чтобы

$$\vec{k}_0 = -\vec{k}_1 \text{ и } |\vec{k}| = \frac{2|\vec{k}_0|}{|\cos(\varphi)|}.\quad (17)$$

Здесь φ — угол между векторами $\vec{k}_0$ и $\vec{k}$. Отметим, что условия Брэгга в этих условиях требуют $|\vec{k}| = 2|\vec{k}_0|$.

4 Взаимодействие волн в нелинейных средах

Уравнения, которые описывают взаимодействие трех волн $(\omega_0, \omega_1, \omega_2; \vec{k}_0, \vec{k}_1, \vec{k}_2)$ в нелинейной среде, можно представить в виде (смотри, например, [2,3])

$$\begin{aligned} \partial a_0 / \partial l_0 &\equiv \dot{a}_0 + (\vec{V}_0 \vec{\nabla}) a_0 = -\sigma_0 \cdot a_1 \cdot a_2 \cdot \exp(-i \cdot \delta); \quad \partial a_1 / \partial l_1 \equiv \dot{a}_1 + (\vec{V}_1 \vec{\nabla}) a_1 = \sigma_1 \cdot a_0 \cdot a_2^* \cdot \exp(i \cdot \delta); \\ \partial a_2 / \partial l_2 &\equiv \dot{a}_2 + (\vec{V}_2 \vec{\nabla}) a_2 = \sigma_2 \cdot a_0 \cdot a_1^* \cdot \exp(i \cdot \delta). \end{aligned} \quad (18)$$

Здесь $\vec{V}_i$ — групповые скорости волн; σ_i — матричные элементы нелинейного взаимодействия.

Левая часть каждого из уравнений системы (18) представлена в виде производных вдоль характеристических направлений. Для определенности будем рассматривать взаимодействие волн с положительной энергией ($\sigma_i > 0$), а также будем ориентироваться на процессы распада. Условие синхронизма взаимодействующих волн будет условие параллельности характеристических прямых гиперплоскости $\delta(\vec{r}, t) = \text{const}$, т.е. условия

$$\Delta\omega - \Delta\vec{k} \cdot \vec{V}_i = 0. \quad (19)$$

Рис. 4. Схема взаимного расположения векторов $\Delta\vec{k}$, $\vec{k}_0$ и $\vec{k}_1$, в стационарной среде при $\Omega = \Delta\omega = 0$.

Рис. 5. Схема распада поперечной волны на поперечную волну и на плазменную волну.

Линейная стадия процесса распада протекает как неустойчивость. На этой стадии амплитуду волны с максимальной частотой можно считать фиксированной ($a_0 = \text{const}$). В этом случае систему (18) удобно переписать в виде

$$\partial a_1 / \partial l_1 = \sigma_1 \cdot a_0 \cdot a_2^* \cdot \exp(i \cdot \delta); \quad \partial a_2^* / \partial l_2 = \sigma_2 \cdot a_0^* \cdot a_1 \cdot \exp(-i \cdot \delta). \quad (20)$$

Подставляя в (20) решение в виде $\sim \exp(i\Omega t - i\vec{\kappa}\vec{r})$, получим следующее дисперсионное уравнение, определяющее связь частоты Ω с вектором $\vec{\kappa}$

$$\Omega^2 - \Omega \cdot \vec{\kappa} \cdot (\vec{V}_1 + \vec{V}_2) + [\kappa^2 \cdot \vec{V}_1 \cdot \vec{V}_2 + \sigma_1 \sigma_2 |a_0|^2] = 0. \quad (21)$$

При решении (21) необходимо из всего множества возможных решений принимать во внимание только те, для которых частота Ω и вектор $\vec{\kappa}$ будут малыми величинами ($\Omega \sim \kappa \sim |a_0|$). Решая уравнение (21) относительно Ω , в этом случае получим

$$\Omega = \vec{\kappa} \vec{V} \pm i \cdot |a_0| \sqrt{\sigma_1 \sigma_2}. \quad (22)$$

Мнимая часть частоты ($\text{Im}\Omega = |a_0| \sqrt{\sigma_1 \sigma_2}$) определяет инкремент распадной неустойчивости.

Представляет интерес найти те новые возможности для распадов, которые не укладываются в известные условия распадов. Рассмотрим простейшие. Пусть происходит распад поперечной волны на попереч-

ную волну и на плазменную волну ($t_0 \rightarrow t_1 + l$). Будем считать, что все три волны укладываются на линейный участок дисперсии (см. рис. 5). В этом случае групповые скорости всех трех волн совпадают ($\vec{V}_0 = \vec{V}_1 = \vec{V}_2 \equiv \vec{V}$). Более того, групповые скорости в этом случае совпадают с фазовыми скоростями. Условия синхронизма (19) в этом случае будут выполняться для любой тройки волн. Действительно, условия (19) приобретают в этом случае вид тождества

$$\omega_0 - \omega_1 - \omega_2 = \left(\frac{\omega_0}{V} - \frac{\omega_1}{V} - \frac{\omega_2}{V} \right) V. \quad (23)$$

Таким образом, в рассматриваемом случае в процессе распада может принимать участие большое количество троек волн. В этом случае процесс распада становится хаотическим. Покажем это. Система уравнений (20) имеет следующий интеграл:

$$\sigma_2 |a_1|^2 - \sigma_1 |a_2|^2 = C = \text{const}. \quad (24)$$

Удобно сделать следующую замену зависимых комплексных переменных $a_j = |a_j| \exp(i\Phi_j)$. Далее учтем, что на начальном этапе процесса распада амплитуды $|a_1|$ и $|a_2|$ малы ($C \rightarrow 0$), а также что $l_0 = l_1 = l_2 \equiv l$ и $(\partial \delta / \partial l) = 0$. В этом случае легко увидеть, что динамика фазы $\Phi = 2(\Phi_0 - \Phi_1 - \Phi_2 + \delta)$ подчиняется уравнению математического маятника

$$\frac{\partial^2 \Phi}{\partial l^2} + \left(2|a_0| \sqrt{\sigma_1 \sigma_2} \right)^2 \sin \Phi = 0. \quad (25)$$

Одновременно с распадом поперечной волны на плазменную волну и на поперечную волну происходит распад этой поперечной волны на плазменную волну и на другую поперечную волну (назовем ее третьей волной), которая лежит на дисперсионной линии (рис. 5) и может находиться сколь угодно близко ко второй поперечной волне. Очевидным является тот факт, что процесс взаимодействия этой новой тройки волн будет описываться системой уравнений аналогичной системе (18). Величина фазы $\Psi = 2(\Phi_0 - \Phi_3 - \Phi_2 + \delta_1)$ при этом будет подчиняться уравнению математического маятника аналогичному (25). Расстояние между нелинейными резонансами при этом может быть сколь угодно малым. Нелинейные резонансы перекрываются. Режим распада становится хаотическим [5,6].

Заметим, что рассматриваемый случай совпадения всех скоростей взаимодействующих волн может быть решен аналитически не только на линейной стадии, но и на нелинейной. Для этого достаточно воспользоваться методом эффективного потенциала (см., например, [2]). Однако в связи с тем, что процесс распада носит хаотический характер, это решение теряет смысл.

5. Взаимодействие в слое

Представляет интерес рассмотреть случай, в котором содержались бы основные элементы рассматриваемых новых условий взаимодействия, и который был бы достаточно простым, чтобы его можно было реализовать в эксперименте. В этом разделе будет рассмотрен такой случай. Он представляет собой слой толщины L неоднородной среды. Границы слоя идеально отражающие. Внутри слоя перпендикулярно его границам распространяются две волны. Каждая из этих волн представляет собой сумму двух волн, одна из которых распространяется вдоль оси z , а вторая — в противоположном направлении (отраженная от границы волна). Динамика взаимодействия каждой из этих волн была описана выше. Выражение для компонент полей в такой системе можно представить в таком виде

$$\begin{aligned} E_x &= [A \exp(ik_0 z) + B \exp(-ik_0 z)] \cdot \exp(-i\omega_0 t) + [C \exp(ik_1 z) + D \exp(-ik_1 z)] \cdot \exp(-i\omega_1 t), \\ H_y &= [A \exp(ik_0 z) - B \exp(-ik_0 z)] \cdot \sqrt{\varepsilon_0} \cdot \exp(-i\omega_0 t) + [C \exp(ik_1 z) - D \exp(-ik_1 z)] \cdot \sqrt{\varepsilon_0} \cdot \exp(-i\omega_1 t). \end{aligned} \quad (26)$$

В формуле (26) $0 < z < L$, $\varepsilon_0 \neq 1$, а A, B, C, D являются медленными функциями времени и координаты.

Аналитический вид этих функций определяется решениями уравнений (8) и (9). Запишем эти решения в следующем виде:

$$A = a \cdot \cos(K\eta); B = b \cdot \cos(K\xi); C = c \cdot \sin(K\xi); D = d \cdot \sin(K\xi). \quad (27)$$

Здесь можно положить $a = -b$, $a = c = d$.

Тогда выражение для электрической компоненты полного поля в слое (26) можно представить в виде

$$\begin{aligned} E_x &= a [\cos(K\eta) \exp(ik_0 z) - \cos(K\xi) \exp(-ik_0 z)] \cdot \exp(-i\omega_0 t) + \\ &+ a [\sin(K\eta) \exp(ik_1 z) + \sin(K\xi) \exp(-ik_1 z)] \cdot \exp(-i\omega_1 t). \end{aligned} \quad (28)$$

Это поле должно обращаться в ноль на поверхности $z = 0$ и на поверхности $z = L$.

Легко найти, что для удовлетворения этому условию волновые вектора k_0 , k_1 и K должны удовлетворять следующим соотношениям:

$$k_0 = \frac{\pi(n+m)}{2L}, \quad k_1 = \frac{\pi(l+p)}{2L}, \quad K = \frac{\pi(n-m)}{2L}. \quad (29)$$

Здесь n, m, l, p — целые числа. Причем, выполняются также условия $(n-m) = (l-p)$, но $(n+m) \neq (l+p)$.

Учитывая, что волновой вектор K пропорционален малому параметру $q \ll 1$ выражения (29) можно упростить. Кроме того, используя выражение для K (8), можно выписать простые выражения для основных параметров рассматриваемой системы

$$\lambda_0 \approx \frac{2L}{n}, \quad n \gg 1; \quad \lambda_1 \approx \frac{2L}{l}, \quad l \gg 1; \quad q \approx \frac{4}{\sqrt{nl}}.$$

Заключение

Таким образом, полученные выше результаты показывают, что действительно, кроме известных условий взаимодействия волн ($\Delta k = \Delta \omega = 0$) существуют более общие условия обмена энергией между волнами. В общем случае, они определяются соотношениями (12). В простом одномерном случае эти условия имеют простой вид $\Delta k = \frac{\Delta \omega \sqrt{\varepsilon}}{c}$. В этих условиях расстройки Δk и $\Delta \omega$ практически произвольные величины. Как видно эти новые условия содержат как частный случай известные условия синхронизма $\Delta k = \Delta \omega = 0$. Наиболее важным во всех этих случаях является тот факт, что расстройка вдоль одной пространственно-временной оси может быть скомпенсирована расстройкой вдоль другой оси. Чтобы наглядно увидеть разницу в старых и новых условиях, на рис. 5 представлен пример спектра волн, которые могут эффективно взаимодействовать. Если использовать известные условия взаимодействия, то для большинства этих волн они выполняться не могут. Используем новые условия синхронизма. Легко увидеть, что любые тройки волн, которые могут быть представлены на рисунке, удовлетворяют новым условиям синхронизма $\Delta k = \Delta \omega \sqrt{\varepsilon} / c$. Они могут эффективно обмениваться энергией. Таких троек неограниченно много.

Полезным, наверное, будет сформулировать аргументы, которые позволяют пренебречь вторыми производными в системах уравнений (5) или (6). Прежде всего, ясно, что это можно сделать из простых физических соображений. В физической литературе это достаточно известный прием. Во-вторых, глядя на уравнения (6), видно, что всегда можно выбрать такие решения этих уравнений, которые зависят только от одной независимой переменной или от ξ или от η . И, наконец, можно воспользоваться методом возмущений для уравнений в частных производных второго порядка, который изложен в монографии [7]. В этом методе анализируется устойчивость скачков, которые распространяются вдоль субхарактеристик. Используя этот метод, можно показать корректность отбрасывания вторых производных, а также корректность полученных выше решений.

Важным примером рассматриваемых условий является пример взаимодействия волн в плоском слое, который был рассмотрен выше. Действительно, как видно из формулы (28), выражение для действительной компоненты электрического поля может быть легко измерено в реальных экспериментах. Видно, что может легко наблюдаться появление волн с новыми частотами (ω_1), если в начальный момент времени в слое существовала только одна волна с частотой (ω_0).

Авторы благодарны Вавриву Д. М. и Неруху А. Г. за полезные дискуссии и замечания.

Список литературы

1. Заславский Г. М., Сагдеев Р. З. Введение в нелинейную физику. М.: Наука, 1988. 368 с.
2. Вильгельмссон Х., Вейланд Я. Когерентное нелинейное взаимодействие волн в плазме. М.: Энергоиздат, 1981. 224 с.
3. Milos, M., Scoric. Nonlinear Physics of Plasmas: Spatiotemporal Structures in Strong Turbulence. Lecture Notes, Research Report NIFS-PROC Series, Toki, Japan, May 2008. 181p.
4. Кадомцев Б. Б. Коллективные явления в плазме. М.: Наука, 1976. 240 с.
5. Буц В. А., Куприянов А. Н., Мануйленко О. В., Толстолужский А. П. // Известия ВУЗов. «ПНД». 1993. Т. 1. № 1. № 2. С. 57–62.
6. Буц В. А., Мануйленко О. В., Степанов К. Н., Толстолужский А. П. // Физика плазмы. 1994. Т. 20. № 9. С. 794–801.
7. Julian, D. Cole. Perturbation Methods in Applied Mathematics, Blaisdell Publishing Company. 1968. Toronto, London. 276 p.

Поступила 23 июня 2020 г.

ABOUT CONDITIONS OF WAVE INTERACTION

V. A. BUTS, D. V. TARASOV

doi: 10.25210/jfop-2002-082090

Based on the three-wave interaction model, new conditions for the interaction of waves are described. These conditions are more general than the known conditions. They, as a special case, contain well-known conditions. New conditions reflect the fact that in distributed systems, the detuning along one of the spatiotemporal axes can be compensated by the detuning along other axes. The examples of the interaction of waves whose characteristics satisfy new conditions are considered. These examples describe the interaction of waves in media with one-dimensional inhomogeneity, in media with three-dimensional inhomogeneity, the interaction of waves in nonlinear media, and also the interaction in a layer. The considered examples make it possible to detect new conditions in the experiment. In particular, when describing the interaction of waves in an inhomogeneous plane layer, expressions for the real component of the electric field are obtained. This characteristic can be measured in real experiments. In addition, it was shown that the appearance of waves with new frequencies (ω_1) can be observed if at the initial moment of time in the layer there was only one wave with frequency (ω_0). This feature of the considered wave interaction is most easily detected in experiments.

Уважаемые читатели!

Подписка на журнал оформляется через подписное агентство «Книга-Сервис» и объединенный каталог «Пресса России» (подписной индекс **29196**)

Журнал включен:

- В *Российский индекс научного цитирования (РИНЦ)*.
- В базу данных *Russian Science Citation Index (RSCI)* на платформе *Web of Science*.
- В *Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук*.

РУБРИКИ ЖУРНАЛА

Обзоры Reviews	Физические основы приборостроения для биологии и медицины Physical Bases of Instrumentation for Biology and Medicine
Математическое моделирование физических процессов Mathematical Modeling of Physical Processes	Методы высокоточных измерений и воспроизведения физических величин Methods of High-Precision Measurements and Reproduction of Physical Values
Приборы и методы быстропротекающих физических процессов Devices and Methods of Fast-Proceeding Physical Processes	Приборы и методы экспериментальной физики The devices and methods of experimental physics
Приборы и методы оценки погрешности физических величин Devices and Methods of an Assessment of an Error of Physical Values	Физические основы приборов и устройств Physical Bases of Instruments and Devices
Приборы и методы физики и техники СВЧ диапазона Microwave Range Devices and Methods of Physics and Techniques	Приборы и устройства Instruments and Devices
Приборы и методы измерений в миллиметровом и субмиллиметровом диапазонах Devices and Methods of Measurements in Millimeter and Submillimeter Ranges	Оптические материалы и технологии Optical Materials and Technologies
Лазерная физика и техника Laser Physics and Technique	Спектральные методы и устройства Spectral Methods and Devices
Приборы и методы дистанционного зондирования Devices and Methods of Remote Sensing	Акустооптические и оптоакустические методы и устройства Acousto-optic and Opto-Acoustic Methods and Devices
Радиолокационные средства и системы Radars and Systems	Физика плазмы в приборостроении Plasma Physics in Instrumentation
Физические основы приборов и устройств метрологии Physical Bases of Metrology Instruments and Devices	Физика и техника плазмы Physics and Technique of Plasma
Проблемы оптического приборостроения на основе метаматериалов Problems of Optical Instrumentation on the Basis of Metamaterials	Новые математические и физические методы New Mathematical and Physical Methods
Проблемы обработки сигналов и изображений в акустооптике и радиофизике Problems of Signal and Image Processing in Acousto-Optics and Radiophysics	R-функции, атомарные функции, вейвлеты, фракталы и хаос в физике и технике R-functions, Atomic Functions, Wavelets, Fractals, and Chaos in Physics and Engineering
Проблемы атомной науки и техники Problems of Atomic Science and Technology	Электродинамика искусственных сред и структур Electrodynamics of Artificial Materials and Structures
Информационно-измерительные системы Information-Measuring System	Электродинамика низко и высокотемпературных сверхпроводящих структур Electrodynamics of Low- and High-Temperature Superconducting Structures
Методы радиолокационных и радиометрических измерений Methods of Radar and Radiometric Measurements	Новости науки Science News
Физические основы космического приборостроения Physical Bases of Space Instrumentation	Краткие сообщения Brief Communications
	Памятные даты Memorable Dates
	Наши юбиляры Our Hero of the Day
	В помощь экспериментатору To Help the Experimenter
	Информация Information

**ПРАВИЛА ОФОРМЛЕНИЯ РУКОПИСЕЙ,
НАПРАВЛЯЕМЫХ В РЕДАКЦИЮ ЖУРНАЛА
«ФИЗИЧЕСКИЕ ОСНОВЫ ПРИБОРОСТРОЕНИЯ»**

1. К публикации принимаются оригинальные работы, не предназначенные для публикации в других изданиях.
2. Необходимы следующие материалы:
 - сопроводительное письмо,
 - акт экспертизы,
 - соответствующий индекс универсальной десятичной классификации (УДК) литературы,
 - название на русском и английском языках,
 - фамилии и инициалы авторов на русском и английском языках (фамилии следуют в алфавитном порядке),
 - аннотацию на русском и английском языках,
 - ключевые слова на русском и английском языках,
 - краткое содержание статьи (реферат) на русском и английском языках объемом 1–2 с,
 - текст статьи вместе с рисунками и таблицами,
 - список литературы в соответствии с ГОСТ Р 7.0.5-2008,
 - краткие сведения об авторах, включающие контактный телефон, фамилию, имя, отчество (полностью), ученую степень (звание), место работы, служебные и домашние адреса авторов с обязательным указанием почтового индекса и номеров телефонов, адрес электронной почты (на отдельной странице на русском и английском языках).
3. Статья представляется в виде файлов формата **MS Word 97-2003 (*.doc)** и одного экземпляра распечатки с полуротным межстрочным интервалом. Файлы могут быть записаны на оптический (CD, DVD) носитель, а также высланы по электронной почте редакции. Статья должна иметь сквозную нумерацию.
4. При наборе текста используются стандартный True Type шрифт — Times New Roman. Устанавливаемый размер бумаги — A4 (210×297 мм).
5. Формульные выражения, а также отдельные символы выполняются исключительно в редакторе формул **MathType**.
6. Иллюстрации выполняются в виде **графических файлов** (цветных/черно-белых) в **растровом** (*.tiff с сжатием без потери качества (LZW) или *.png) или **векторном** формате. Разрешение изображений должно быть не хуже **300 dpi** и шириной не менее 8 см (абсолютные размеры изображения **не менее 1000×1000 точек**). Все иллюстрации сопровождаются подрисовочными подписями (не повторяющими фразы-ссылки на рисунки в тексте), включающими в себя номер, название и при необходимости условные обозначения. Рисунки должны соответствовать следующим требованиям:
 - буквенные и цифровые обозначения на рисунках по **начертанию** и **размеру** должны соответствовать обозначениям в тексте статьи,
 - текстовая информация и условные обозначения выносятся из рисунка в текст статьи или подрисовочные подписи,
 - для оформления рисунков **не допускается применять средства рисования Word**, а также таблицы. Режим обтекания только «в тексте» с выравниванием по центру.Иллюстрации должны быть включены в текст, а также представлены отдельными графическими файлами.
7. Термины, определения и единицы физических величин должны соответствовать действующим ГОСТ.
8. В формулах латинские буквы следует набирать курсивом, а цифры, греческие строчные и прописные символы прямо. Векторы и матрицы следует набирать прямыми жирным шрифтом; «exp(.)» в значении экспоненты набирать прямым светлым шрифтом. В индексах сокращения от русских и английских слов следует набирать прямым шрифтом.
9. Формулы нумеруются в круглых скобках (например, (2)), литературные ссылки в прямых — [2], подстрочные замечания оформляются сноской.
10. На последней странице рукописи должны быть подписи всех авторов.
11. Редакция не ставит в известность авторов об изменениях и сокращениях рукописи, имеющих редакционный характер и не затрагивающих принципиальных вопросов.
12. Рукописи, в которых не соблюдены данные требования, возвращаются авторам без рассмотрения.
13. Редакция оставляет за собой право использовать наиболее интересные статьи в периодических тематических сборниках библиотек журнала. Предоставление редакции рукописи является подтверждением согласия автора на указанное выше использование его произведения.
14. Авторы статей несут ответственность за полноту и достоверность цитируемой в них литературы, а также за публикацию заимствованного материала без ссылки на источник.
15. За достоверность сведений, изложенных в публикациях, редакция и издатель ответственности не несут.
16. За публикацию материалов, содержащих закрытые сведения, авторы несут персональную ответственность на основании действующих законодательных актов.
17. Итоговое решение об одобрении или отклонении представленного в редакцию материала принимается редакционной коллегией и является окончательным.

**Уважаемые коллеги, приглашаем вас принять участие
в 13-ой Международной конференции
«Акустооптические и радиолокационные методы
измерений и обработки информации»**

Время и место проведения: 12 — 13 октября 2020,
Научно-технологический центр уникального приборостроения
Российской академии наук, Москва, ул. Бутлерова д. 15

E-mail: expo@ntcup.ru

WEB: <http://armimp.ru/a2020/>

ОРГАНИЗАТОРЫ:

Научно-технологический центр уникального приборостроения РАН,
Институт радиотехники и электроники им. В.А. Котельникова РАН,
Московский государственный технический университет им. Н.Э. Баумана,
Московский государственный университет им. М.В. Ломоносова,
РНТОРЭС им. А.С. Попова

Программа конференции:

ПЛЕНАРНОЕ ЗАСЕДАНИЕ

СЕКЦИЯ 1. Методы математического моделирования физических процессов
в оптике и радиолокации. R-функции, атомарные функции, вейвлеты, фракталы и хаос

СЕКЦИЯ 2. Генерирование, излучение и распространение сверхширокополосных сигналов
и сверхкоротких импульсов

СЕКЦИЯ 3. Физические основы приборостроения, акустооптические методы и устройства,
фурье-спектроскопия

Основные научные направления:

- генерирование и излучение сверхширокополосных сигналов и сверхкоротких импульсов;
- прием, измерение и обработка сверхширокополосных сигналов и сверхкоротких импульсов;
- распространение сверхширокополосных сигналов и сверхкоротких импульсов в природных средах;
- зондирование природных сред сверхширокополосными сигналами и сверхкороткими импульсами;
- миллиметровые и субмиллиметровые волны;
- методы математического моделирования физических процессов в оптике и радиолокации;
- методы решения некорректных задач в радиолокации и оптике;
- R-функции, атомарные функции, вейвлеты, фракталы и хаос;
- информационно-измерительные оптические и радиотехнические системы;
- сверхширокополосные хаотические сигналы в оптических и радиотехнических информационных системах;
- методы вычислительной томографии в оптике и радиолокации;
- лазерная физика и техника;
- сверхкороткие импульсы и методы нелинейной оптики;
- акустооптические и радиолокационные методы измерений и обработки в биологии, медицине;
- физические основы приборостроения;
- космическое приборостроение;
- методы неразрушающего контроля;
- биомедицинская информатика.

Работа конференции будет проходить в залах, оборудованных мультимедийным проектором.
Формат презентаций **Power Point** и **Adobe PDF**.

ARMIMP-2020

<http://armimp.ru/a2020/>

ИНФОРМАЦИЯ

УВАЖАЕМЫЕ ЧИТАТЕЛИ, В ИЗДАТЕЛЬСТВЕ «ТЕХНОСФЕРА» ВЫШЛИ МОНОГРАФИИ

Кравченко В.Ф., Чуриков Д.В. Цифровая обработка сигналов атомарными функциями и вейвлетами. М.: Техносфера, 2019 (Дополнительный тираж). 182с.

ISBN 978-5-94836-506-0

Под редакцией заслуженного деятеля науки РФ,
доктора физико-математических наук, профессора В.Ф. Кравченко

Аннотация

Монография посвящена исследованию, развитию и разработке методов анализа физических систем на основе теорий атомарных, R- и WA- систем функций, построению алгоритмов обработки и моделирования процессов дистанционного зондирования и радиоастрономии с целью улучшения физических характеристик передачи, восстановления и распознавания многомерных сигналов. Рассматриваются новые весовые функции (окна) и фильтры с конечной импульсной характеристикой на их основе, а также ортогональные и аналитические вейвлеты. Исследована обобщенная теорема отсчетов на основе атомарных функций и её частные случаи. Разработан алгоритм синтеза многомерных функций с произвольной геометрией опорной области. Предложено и обосновано обобщение функции неопределенности по времени и частоте применительно к анализу сложных зондирующих сигналов.

Книга предназначена для научных работников, аспирантов и студентов, работающих в области цифровой обработки сигналов применительно к современным задачам радиопизики и радиотехники.

Кравченко В.Ф., Кравченко О.В. Конструктивные методы алгебры логики, атомарных функций, вейвлетов, фракталов в задачах физики и техники. М.: Техносфера, 2018. 696с.

ISBN 978-5-94836-518-3

Под редакцией заслуженного деятеля науки РФ,
доктора физико-математических наук, профессора В.Ф. Кравченко

Аннотация

В монографии изложены конструктивные методы алгебры логики, теория R-функций (функций В.Л. Рвачева), атомарных функций (АФ), вейвлетов, WA-систем функций и фракталов применительно к задачам современной физики и техники. Теория R-функций, содержащая в себе идеи булевой алгебры и аналитической геометрии, представляет собой универсальный математический аппарат, позволяющий в неявном виде получать уравнения границ областей произвольной геометрии. На основе этих уравнений осуществляется построение структур решения краевых задач в областях сложной формы, т.е. таких аналитических выражений, которые априори удовлетворяют заданным краевым условиям. Структуры решения зависят от неопределенных компонент, находящихся из условия минимизации функционала Ритца или одним из проекционных методов.

В главах 1–4 и приложениях описаны алгоритмы и приведены многочисленные примеры решения различных типов внешних и внутренних краевых задач электростатики и электродинамики в областях сложной формы. В главе 5 рассмотрено построение новых ортогональных вейвлетов Кравченко на основе атомарных функций. Глава 6 посвящена построению новых конструкций ортогональных вейвлетов на основе АФ $h(x)$. Обосновано обобщение функции неопределенности (ФН) по времени и частоте на основе семейства АФ применительно к цифровой обработке сигналов в антенных системах. Исследован новый класс аналитических вейвлетов Кравченко–Рвачева (АКР-вейвлетов). В главе 7 предложен и обоснован новый метод решения краевых задач для дифференциальных уравнений в частных производных эллиптического типа на основе R-функций и вейвлетов применительно к волноводам сложного поперечного сечения. В главе 8 и приложении к ней построен и обоснован новый класс фрактальных функций с учетом специфических свойств атомарных функций. Впервые получены новые синтезированные атомарно-фрактальные функции в комбинациях с классическими фрактальными функциями. Проведен численный эксперимент для одномерных, двумерных и кольцевых фрактальных антенных решеток.

Монография представляет интерес для научных работников, аспирантов и студентов старших курсов радиофизических и радиотехнических специальностей, а также специалистов, работающих в области вычислительной математики и физики. Рукопись монографии Кравченко В.Ф., Кравченко О.В. «Конструктивные методы алгебры логики, атомарных функций, вейвлетов, фракталов в задачах физики и техники» была рассмотрена на заседании Ученого совета Института радиотехники и электроники им. В.А. Котельникова РАН 25 декабря 2015 года (выписка из протокола № 7) и рекомендована к опубликованию. При оформлении обложки использована иллюстрация знаменитой шахматной партии (Р. Фишер — М. Тайманов, Ванкувер, 1971), комментарий см. на стр. 695.

мир цифровой обработки

В.Ф. КРАВЧЕНКО,
Д.В. ЧУРИКОВ

Цифровая обработка
сигналов атомарными
функциями и вейвлетами

КРАВЧЕНКО ВИКТОР ФИЛИППОВИЧ –
заслуженный деятель науки РФ, д.ф.-м.н.,
профессор, институт радиотехники и
электроники им. В. А. Котельникова РАН,
научно-технологический центр
уникального приборостроения РАН,
Московский государственный
технический университет им. Н.Э. Баумана

ЧУРИКОВ ДМИТРИЙ ВИКТОРОВИЧ –
к.ф.-м.н., институт радиотехники
и электроники им. В. А. Котельникова РАН,
научно-технологический центр
уникального приборостроения РАН,
Московский физико-технический
институт (государственный университет)
РОССИЙСКИЙ НОВЫЙ УНИВЕРСИТЕТ

- ИССЛЕДОВАНИЕ, РАЗВИТИЕ И РАЗРАБОТКА
МЕТОДОВ АНАЛИЗА ФИЗИЧЕСКИХ СИСТЕМ
НА ОСНОВЕ ТЕОРИИ АТОМАРНЫХ,
• R- И WA-СИСТЕМ ФУНКЦИЙ
• ПОСТРОЕНИЕ АЛГОРИТМОВ ОБРАБОТКИ И
МОДЕЛИРОВАНИЯ ПРОЦЕССОВ
ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ И
РАДИОАСТРОНОМИИ В ЦЕЛЯХ УЛУЧШЕНИЯ
ФИЗИЧЕСКИХ ХАРАКТЕРИСТИК ПЕРЕДАЧИ
• ВОССТАНОВЛЕНИЯ И РАСПОЗНАВАНИЯ
МНОГОМЕРНЫХ СИГНАЛОВ.

ТЕХНОСФЕРА

Цифровая обработка
сигналов атомарными
функциями и вейвлетами

цифровой обработки

XI
♦
15

цифровой обработки

В.Ф. КРАВЧЕНКО,
Д.В. ЧУРИКОВ

Цифровая обработка
сигналов атомарными
функциями и вейвлетами

ТЕХНОСФЕРА

Харьковский национальный университет
им. В.Н. Каразина

Московский государственный
технический университет
им. Н.Э. Баумана

Институт радиотехники и электроники
им. В.А. Котельникова РАН

В.Ф. Кравченко
О.В. Кравченко

КОНСТРУКТИВНЫЕ МЕТОДЫ АЛГЕБРЫ
ЛОГИКИ, АТОМАРНЫХ ФУНКЦИЙ,
ВЕЙВЛЕТОВ, ФРАКТАЛОВ В ЗАДАЧАХ
ФИЗИКИ И ТЕХНИКИ

В.Ф. Кравченко, О.В. Кравченко

КОНСТРУКТИВНЫЕ МЕТОДЫ АЛГЕБРЫ
ЛОГИКИ, АТОМАРНЫХ ФУНКЦИЙ,
ВЕЙВЛЕТОВ, ФРАКТАЛОВ В ЗАДАЧАХ
ФИЗИКИ И ТЕХНИКИ

ТЕХНОСФЕРА

ISSN 2225-4293

9 772225 429003 >

<http://jfop.ru>

<http://dx.doi.org/10.25210/jfop-2002>