

ISSN 1816-0301

ИНФОРМАТИКА

4(52)

ОКТЯБРЬ-ДЕКАБРЬ
2016

Редакционная коллегия:

Главный редактор

А.В. Тузиков

Заместитель главного редактора

М.Я. Ковалев

Члены редколлегии

С.В. Абламейко, В.В. Анищенко, П.Н. Бибило, М.Н. Бобов,
А.Н. Дудин, С.Я. Килин, В.В. Краснопрошин, А.М. Крот, С.В. Кругликов,
С.П. Кундас, Н.А. Лиходед, П.П. Матус, С.В. Медведев, А.А. Петровский,
Ю.Н. Сотсков, Ю.С. Харин, А.Ф. Чернявский, В.Н. Ярмолик
Н.А. Рудая (*заведующая редакцией*)

Адрес редакции:

220012, Минск,
ул. Сурганова, 6, к. 305
тел. (017) 284-26-22
e-mail: rio@newman.bas-net.by
<http://uiip.bas-net.by>

ОБЪЕДИНЕННЫЙ ИНСТИТУТ ПРОБЛЕМ ИНФОРМАТИКИ
НАЦИОНАЛЬНОЙ АКАДЕМИИ НАУК БЕЛАРУСИ

ИНФОРМАТИКА

ЕЖЕКВАРТАЛЬНЫЙ НАУЧНЫЙ ЖУРНАЛ

Издается с января 2004 г.

№ 4(52) • октябрь-декабрь 2016

СОДЕРЖАНИЕ

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

- Ерофеев В.Т.** Многократная фокусировка электромагнитного поля магнитного диполя плоскопараллельной двухслойной линзой из метаматериала5
- Медвидь Н.В., Мартынюк П.Н.** Реализация алгоритма бессеточного метода радиальных базисных функций в задачах подземной гидромеханики20
- Курочка К.С.** Конечно-элементное моделирование распределения электромагнитного поля в ближней зоне сферической наночастицы33

ОБРАБОТКА СИГНАЛОВ, ИЗОБРАЖЕНИЙ И РЕЧИ

- Борискевич А.А.** Внедрение идентификационных данных в изображения на основе синусоидальных решеток для мобильных приложений42
- Артемьев В.М., Наумов А.О., Кохан Л.Л.** Линейная фильтрация многомерных случайных последовательностей расширенным методом наименьших квадратов51

К 80-ЛЕТИЮ ЧЛЕНА-КОРРЕСПОНДЕНТА Г.В. РИМСКОГО

Вишняков В.А. Научное наследие члена-корреспондента Г.В. Римского.....	57
Несенчук А.А. Исследование поведения и синтез семейств характеристических полиномов интервальных динамических систем на основе использования функции параметра корневого портрета системы.....	61
Ткаченко В.В. Применение методов общей теории корневых траекторий в исследовании систем многосвязного регулирования	74
Романюк Г.Э., Ходар М.С. Анализ лексикографической окрашенности художественных текстов	82
Серебряная Л.В., Потараев В.В. Методы классификации текстовой информации на основе искусственной нейронной и семантической сетей	95
Шатохин И.В. Информационная система управления предприятием как цепочкой поставок.....	104
Смольников М.А., Скудняков Ю.А. Использование агрегации данных в мобильных сенсорных сетях	111

Редактор Г.Б. Гончаренко
Корректор А.А. Михайлова
Компьютерная верстка О.Б. Бутевич

Сдано в набор 03.10.2016. Подписано в печать 21.11.2016.
Формат 60×84 1/8. Бумага офсетная. Гарнитура Таймс. Ризография.
Усл. печ. л. 13,5. Уч.-изд. л. 13,2. Тираж 60 экз. Заказ 9.

Государственное научное учреждение «Объединенный институт проблем информатики Национальной академии наук Беларуси».
Свидетельство о государственной регистрации издателя, изготовителя, распространителя печатных изданий № 1/274 от 04.04.2014.
ЛП № 02330/444 от 18.12.13.
Ул. Сурганова, 6, 220012, Минск.

THE UNITED INSTITUTE OF INFORMATICS PROBLEMS
OF THE NATIONAL ACADEMY OF SCIENCES OF BELARUS

INFORMATICS

PUBLISHED QUATERLY

Issued since 2004

№ 4(52) • October-December 2016

CONTENTS

MATHEMATICAL MODELING

- Erofeenko V.T.** Multiple focusing of electromagnetic field of magnetic dipole by means of plane-parallel two-layer lens from metamaterials5
- Medvid N.V., Martinyuk P.M.** The algorithm of meshfree method of radial basis functions in tasks of underground hydromechanics20
- Kurachka K.S.** Finite-element modeling of electromagnetic field distribution in the nearest zone of a spherical nanoparticle33

SIGNAL, IMAGE AND SPEECH PROCESSING

- Boriskevich A.A.** Embedding identification data within images based on sinusoidal gratings for mobile applications42
- Artemiev V.M., Naumov A.O., Kokhan L.L.** Linear filtration of random sequences using a least square method with regularization51

IN MEMORY OF 80th ANNIVERSARY OF G.V. RIMSKIY

Vishniakou U.A. Scientific legacy of professor G.V. Rimskiy	57
Nesenchuk A.A. Investigation of the interval dynamic systems behavior and synthesis of their characteristic polynomials on the basis of the system root locus portraits parameter function application	61
Tkachenko V.V. The application of general theory of root trajectories methods in research of multi-channel control system with cross-connections	74
Romaniuk G.E., Khodar M.S. The analysis of lexicographical coloring the literary texts	82
Serebryanaya L.V., Potaraev V.V. Methods of text information classification on the basis of artificial neural and semantic networks.....	95
Shatokhin I.V. Information system of enterprise supply chain management.....	104
Smolnikov M.A., Skudnyakov Y.A. The use of data aggregation in mobile sensor networks	111

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

УДК 537.874:621.372.8

В.Т. Ерофеев

**МНОГОКРАТНАЯ ФОКУСИРОВКА ЭЛЕКТРОМАГНИТНОГО ПОЛЯ
МАГНИТНОГО ДИПОЛЯ ПЛОСКОПАРАЛЛЕЛЬНОЙ ДВУХСЛОЙНОЙ
ЛИНЗОЙ ИЗ МЕТАМАТЕРИАЛА**

Предлагается метод аналитического решения краевой задачи проникновения электромагнитного поля точечного магнитного диполя через бесконечную плоскую двухслойную линзу (экран) из фокусирующих метаматериалов. Показывается, что для линз со специальными комбинациями материальных параметров слоев образуется серия фокусов: один за экраном и три перед экраном в полупространстве расположения точечного источника поля возбуждения. При этом суммарное поле фокусов перед линзой равно сумме полей магнитных и электрических диполей.

Введение

В электродинамике экранирующих структур значительное место занимает исследование свойств многослойных экранов, слои которых выполнены из различных композитных материалов: бианизотропных, биизотропных, киральных, метаматериалов и др. [1–3]. Как правило, изучаются структуры геометрически правильной формы: сферические, цилиндрические и плоские [4]. Это связано с возможностью получения аналитических решений задач рассеяния электромагнитных волн на экранах [5] и с их практическим применением. Наиболее детально исследованы многослойные плоские экраны при воздействии плоских электромагнитных волн [1, 2, 6]. Разрабатываются методы моделирования, основанные на использовании передаточных матриц и двухсторонних граничных условий для многослойных структур [7–10]. В последнее время в литературе выделяется класс экранов из метаматериалов с фокусирующими свойствами. Создаются модели, описывающие взаимодействие электромагнитных полей точечных источников с экранами из метаматериалов конечных размеров [11, 12] и бесконечно протяженных [13]. Выявляются условия фокусировки точечных источников однослойными и двухслойными экранами со специальными материальными параметрами [14, 15]. Надо отметить, что для экрана с фокусирующими параметрами решение задачи рассеяния поля точечного источника в классическом смысле не существует [16], но в предельном смысле, при приближении материальных параметров экрана к фокусирующим, решение задачи приближается к решению с особыми точками в фокусах [15].

В настоящей работе исследовано интегральное решение краевой задачи прохождения электромагнитного поля магнитного диполя через плоскопараллельную двухслойную линзу (экран) со слоями из метаматериалов. Применена методика моделирования, основанная на использовании двухсторонних граничных условий на поверхности многослойного биизотропного экрана [7] и теорем сложения, связывающих базисные сферические и цилиндрические электромагнитные поля [5, с. 153]. Исследована структура полей за экраном и перед экраном в случае фокусирующих материальных параметров слоев.

1. Постановка задачи

В свободном пространстве R^3 с декартовой системой координат $Oxyz$ разместим плоский экран $D(0 < z < \Delta)$, состоящий из двух слоев $\Omega_1(0 < z < \Delta_1)$, $\Omega_2(\Delta_1 < z < \Delta_1 + \Delta_2)$ с толщинами Δ_s , $s = 1, 2$ ($\Delta_1 + \Delta_2 = \Delta$). Слои выполнены из биизотропных материалов, характеризуемых следующими материальными комплексными параметрами: $\epsilon_s = \epsilon_{rs}\epsilon_0$, $\mu_s = \mu_{rs}\mu_0$,

$G_s = G_{rs}/c$, $Z_s = Z_{rs}/c$; ε_0, μ_0 – электрическая и магнитная постоянные, c – скорость света в вакууме. В полупространстве $D_1(z < 0)$ на оси Oz в точке $O_1(0, 0, -h)$, $h > 0$, размещен точечный магнитный диполь, который колеблется с круговой частотой $\omega = 2\pi f$ и возбуждает первичное поле \vec{E}_0, \vec{H}_0 . В результате взаимодействия поля диполя с экраном D образуются поля: \vec{E}'_1, \vec{H}'_1 – поле в D_1 , отраженное от экрана D ; $\vec{E}_1 = \vec{E}_0 + \vec{E}'_1$, $\vec{H}_1 = \vec{H}_0 + \vec{H}'_1$ – суммарное поле в области D_1 ; $\vec{E}^{(s)}, \vec{H}^{(s)}$ – поле в слое Ω_s ; \vec{E}_2, \vec{H}_2 – поле, проникшее в полупространство $D_2(z > \Delta)$. Рассматривается временная зависимость полей вида $\exp(-i\omega t)$. Сформулируем краевую задачу, моделирующую проникновение первичного поля \vec{E}_0, \vec{H}_0 через экран D .

Краевая задача 1. При заданном поле \vec{E}_0, \vec{H}_0 требуется определить поля \vec{E}'_1, \vec{H}'_1 ; $\vec{E}^{(s)}, \vec{H}^{(s)}$; \vec{E}_2, \vec{H}_2 , которые удовлетворяют уравнениям

$$\text{rot } \vec{E}_j = i\omega\mu_0\vec{H}_j, \text{rot } \vec{H}_j = -i\omega\varepsilon_0\vec{E}_j \text{ в } D_j, \quad (1)$$

$$\text{rot } \vec{E}^{(s)} = i\omega(\mu_s\vec{H}^{(s)} + Z_s\vec{E}^{(s)}), \text{rot } \vec{H}^{(s)} = -i\omega(\varepsilon_s\vec{E}^{(s)} + G_s\vec{H}^{(s)}) \text{ в } \Omega_s,$$

граничным условиям непрерывности тангенциальных составляющих полей на плоскостях раздела сред $\Gamma_1(z = 0)$, $\gamma_1(z = \Delta_1)$, $\Gamma_2(z = \Delta)$

$$\begin{aligned} (\vec{E}_{1r} - \vec{E}_r^{(1)})|_{z=0} = 0, (\vec{H}_{1r} - \vec{H}_r^{(1)})|_{z=0} = 0, (\vec{E}_{2r} - \vec{E}_r^{(2)})|_{z=\Delta} = 0, (\vec{H}_{2r} - \vec{H}_r^{(2)})|_{z=\Delta} = 0, \\ (\vec{E}_r^{(1)} - \vec{E}_r^{(2)})|_{z=\Delta_1} = 0, (\vec{H}_r^{(1)} - \vec{H}_r^{(2)})|_{z=\Delta_1} = 0 \end{aligned} \quad (2)$$

и условиям излучения на бесконечности

$$\lim_{r \rightarrow \infty} r \left(\frac{\partial \vec{E}'_1}{\partial r} - ik_0\vec{E}'_1 \right) = 0, \quad z < 0; \quad \lim_{r \rightarrow \infty} r \left(\frac{\partial \vec{E}_2}{\partial r} - ik_0\vec{E}_2 \right) = 0, \quad z > \Delta, \quad (3)$$

где $k_0 = \frac{\omega}{c}$; r – радиальная координата сферической системы координат $O r \theta \varphi$, связанной с декартовой системой координат $Oxyz$.

2. Интегральное представление полей диполей

Рассмотрим магнитный и электрический диполи, расположенные в начале координат O и ориентированные вдоль оси Oz :

– магнитный диполь (TE -диполь, электрический вектор параллелен экрану D) [5, с. 170]

$$\vec{E}_m(O) = A_0 \tilde{m}_{01}(\vec{r}, k_0), \vec{H}_m(O) = h_0 A_0 \tilde{n}_{01}(\vec{r}, k_0); \quad (4)$$

– электрический диполь (TH -диполь, магнитный вектор параллелен экрану D) [5, с. 162]

$$\vec{E}_e(O) = B_0 \tilde{n}_{01}(\vec{r}, k_0), \vec{H}_e(O) = h_0 B_0 \tilde{m}_{01}(\vec{r}, k_0), \quad (5)$$

где A_0, B_0 – амплитуды диполей;

$$\vec{r} = (r, \theta, \varphi), \quad h_0 = \frac{1}{iZ_0}, \quad Z_0 = \sqrt{\frac{\mu_0}{\varepsilon_0}}, \quad \tilde{m}_{01}, \tilde{n}_{01} \text{ – базисные сферические поля.}$$

Поля точечных диполей, расположенных в вакууме, выражаются через базисные сферические электромагнитные поля [5, с. 118].

Учитывая осевую симметрию полей диполей, представим поля (4), (5) через базисные цилиндрические поля в вакууме [5, с. 130]:

$$\vec{M}_0^{(\mp 1)}(\vec{\rho}; \lambda, k_0) = \vec{V}_0^{(1)}(\lambda \rho) e^{\mp v_0(\lambda)z}, \quad (6)$$

$$\vec{M}_0^{(\mp 2)}(\vec{\rho}; \lambda, k_0) = \frac{1}{k_0} \left(\mp v_0(\lambda) \vec{V}_0^{(2)}(\lambda \rho) + \lambda J_0(\lambda \rho) \vec{e}_z \right) e^{\mp v_0(\lambda)z},$$

где

$$\vec{V}_0^{(1)}(\lambda \rho) = \vec{V}_1 = J_1(\lambda \rho) \vec{e}_\varphi, \quad \vec{V}_0^{(2)}(\lambda \rho) = \vec{V}_2 = -J_1(\lambda \rho) \vec{e}_\rho, \quad (7)$$

$v_0(\lambda) = \sqrt{\lambda^2 - k_0^2}$ при $\lambda \geq k_0$, $v_0(\lambda) = -i\sqrt{k_0^2 - \lambda^2}$ при $k_0 > \lambda$, $0 \leq \lambda < \infty$; $\vec{e}_\rho, \vec{e}_\varphi, \vec{e}_z$ – орты цилиндрической системы координат $O\rho\varphi z$, $\vec{\rho} = (\rho, \varphi, z)$; $J_m(x)$ – функция Бесселя.

Используя теоремы сложения [5, с.153], представим поля диполей (4), (5) через цилиндрические поля (6) в виде интегралов:

$$\begin{aligned} \vec{E}_m(O) &= A_0 \begin{cases} \int_0^\infty c(\lambda) \vec{M}_0^{(-1)}(\vec{\rho}; \lambda, k_0) d\lambda, & z > 0, \\ 0 \\ \int_0^\infty c(\lambda) \vec{M}_0^{(+1)}(\vec{\rho}; \lambda, k_0) d\lambda, & z < 0; \end{cases} \\ \vec{H}_m(O) &= h_0 A_0 \begin{cases} \int_0^\infty c(\lambda) \vec{M}_0^{(-2)}(\vec{\rho}; \lambda, k_0) d\lambda, & z > 0, \\ 0 \\ \int_0^\infty c(\lambda) \vec{M}_0^{(+2)}(\vec{\rho}; \lambda, k_0) d\lambda, & z < 0; \end{cases} \\ \vec{E}_s(O) &= B_0 \begin{cases} \int_0^\infty c(\lambda) \vec{M}_0^{(-2)}(\vec{\rho}; \lambda, k_0) d\lambda, & z > 0, \\ 0 \\ \int_0^\infty c(\lambda) \vec{M}_0^{(+2)}(\vec{\rho}; \lambda, k_0) d\lambda, & z < 0, \end{cases} \\ \vec{H}_s(O) &= h_0 B_0 \begin{cases} \int_0^\infty c(\lambda) \vec{M}_0^{(-1)}(\vec{\rho}; \lambda, k_0) d\lambda, & z > 0, \\ 0 \\ \int_0^\infty c(\lambda) \vec{M}_0^{(+1)}(\vec{\rho}; \lambda, k_0) d\lambda, & z < 0, \end{cases} \end{aligned} \quad (8)$$

где $c(\lambda) = -\frac{i\lambda^2}{k_0^2 v_0(\lambda)}$.

Для диполей (4), (5), расположенных в произвольных точках $O_l(0, 0, h_l)$, $l = 1, 2, 3, \dots$, на оси Oz , учитывая сдвиг системы координат $z = z_l + h_l$ в формулах (8), получим интегральные представления полей:

$$\vec{\mathbf{E}}_{\mathbf{M}}(O_l) = \begin{cases} \int_0^{\infty} A(\lambda, h_l) \vec{M}_0^{(-1)}(\vec{\rho}; \lambda, k_0) d\lambda, & z > h_l, \\ \int_0^{\infty} A(\lambda, -h_l) \vec{M}_0^{(+1)}(\vec{\rho}; \lambda, k_0) d\lambda, & z < h_l, \end{cases}$$

$$\vec{\mathbf{H}}_{\mathbf{M}}(O_l) = h_0 \begin{cases} \int_0^{\infty} A(\lambda, h_l) \vec{M}_0^{(-2)}(\vec{\rho}; \lambda, k_0) d\lambda, & z > h_l, \\ \int_0^{\infty} A(\lambda, -h_l) \vec{M}_0^{(+2)}(\vec{\rho}; \lambda, k_0) d\lambda, & z < h_l; \end{cases} \quad (9)$$

$$\vec{\mathbf{E}}_{\mathbf{s}}(O_l) = \begin{cases} \int_0^{\infty} B(\lambda, h_l) \vec{M}_0^{(-2)}(\vec{\rho}; \lambda, k_0) d\lambda, & z > h_l, \\ \int_0^{\infty} B(\lambda, -h_l) \vec{M}_0^{(+2)}(\vec{\rho}; \lambda, k_0) d\lambda, & z < h_l, \end{cases}$$

$$\vec{\mathbf{H}}_{\mathbf{s}}(O_l) = h_0 \begin{cases} \int_0^{\infty} B(\lambda, h_l) \vec{M}_0^{(-1)}(\vec{\rho}; \lambda, k_0) d\lambda, & z > h_l, \\ \int_0^{\infty} B(\lambda, -h_l) \vec{M}_0^{(+1)}(\vec{\rho}; \lambda, k_0) d\lambda, & z < h_l, \end{cases}$$

где $A(\lambda, h_l) = A_0 c(\lambda) e^{v_0(\lambda)h_l}$, $B(\lambda, h_l) = B_0 c(\lambda) e^{v_0(\lambda)h_l}$. (10)

Для поля магнитного диполя, расположенного в точке $O_{\text{и}}(0, 0, h_{\text{и}}) \in D_1$, $h_{\text{и}} = -h$, из формул (9) следует интегральное представление

$$\vec{\mathbf{E}}_0 = \vec{\mathbf{E}}_{\mathbf{M}}(O_{\text{и}}) = \int_0^{\infty} A(\lambda, -h) \vec{M}_0^{(-1)}(\vec{\rho}; \lambda, k_0) d\lambda, \quad z > -h,$$

$$\vec{\mathbf{H}}_0 = \vec{\mathbf{H}}_{\mathbf{M}}(O_{\text{и}}) = h_0 \int_0^{\infty} A(\lambda, -h) \vec{M}_0^{(-2)}(\vec{\rho}; \lambda, k_0) d\lambda. \quad (11)$$

3. Краевая задача для экрана с двухсторонними граничными условиями

Учитывая структуру первичного поля (11), решение задачи (1)–(3) представим в интегральном виде:

$$\vec{\mathbf{E}}_1' = \int_0^{\infty} \vec{E}_1'(\lambda) d\lambda, \quad \vec{\mathbf{H}}_1' = \int_0^{\infty} \vec{H}_1'(\lambda) d\lambda, \quad z < 0,$$

$$\vec{\mathbf{E}}_2 = \int_0^{\infty} \vec{E}_2(\lambda) d\lambda, \quad \vec{\mathbf{H}}_2 = \int_0^{\infty} \vec{H}_2(\lambda) d\lambda, \quad z > \Delta,$$

$$\vec{\mathbf{E}}^{(s)} = \int_0^{\infty} \vec{E}^s(\lambda) d\lambda, \quad \vec{\mathbf{H}}^{(s)} = \int_0^{\infty} \vec{H}^s(\lambda) d\lambda \quad \text{в } \Omega_s,$$

$$\vec{\mathbf{E}}_0 = \int_0^{\infty} \vec{E}_0(\lambda) d\lambda, \quad \vec{\mathbf{H}}_0 = \int_0^{\infty} \vec{H}_0(\lambda) d\lambda,$$
(12)

где

$$\begin{aligned} \vec{E}_0 + \vec{E}'_1 &= \vec{E}_1, \quad \vec{H}_0 + \vec{H}'_1 = \vec{H}_1, \\ \vec{E}_0(\lambda) &= a(\lambda)\vec{M}_0^{(-1)} + b(\lambda)\vec{M}_0^{(-2)}, \quad \vec{H}_0(\lambda) = h_0(a(\lambda)\vec{M}_0^{(-2)} + b(\lambda)\vec{M}_0^{(-1)}), \end{aligned} \quad (13)$$

$$a(\lambda) = A(\lambda, -h), \quad b(\lambda) = 0.$$

Поля (12) удовлетворяют условиям (1)–(3) краевой задачи 1, если для подынтегральных полей выполнены условия следующей задачи.

Краевая задача 2. При заданном поле (13) требуется определить поля $\vec{E}'_1, \vec{H}'_1; \vec{E}_2, \vec{H}_2; \vec{E}^s, \vec{H}^s$, которые удовлетворяют уравнениям

$$\begin{aligned} \text{rot } \vec{E}_j &= i\omega\mu_0\vec{H}_j, \quad \text{rot } \vec{H}_j = -i\omega\varepsilon_0\vec{E}_j \quad \text{в } D_j, \\ \text{rot } \vec{E}^s &= i\omega(\mu_s\vec{H}^s + Z_s\vec{E}^s), \quad \text{rot } \vec{H}^s = -i\omega(\varepsilon_s\vec{E}^s + G_s\vec{H}^s) \quad \text{в } \Omega_s; \end{aligned} \quad (14)$$

граничным условиям

$$\begin{aligned} (\vec{E}_{1\tau} - \vec{E}'_{1\tau})\Big|_{z=0} &= 0, \quad (\vec{H}_{1\tau} - \vec{H}'_{1\tau})\Big|_{z=0} = 0, \quad (\vec{E}_{2\tau} - \vec{E}'_{2\tau})\Big|_{z=\Delta} = 0, \quad (\vec{H}_{2\tau} - \vec{H}'_{2\tau})\Big|_{z=\Delta} = 0, \\ (\vec{E}_{\tau}^1 - \vec{E}_{\tau}^2)\Big|_{z=\Delta_1} &= 0, \quad (\vec{H}_{\tau}^1 - \vec{H}_{\tau}^2)\Big|_{z=\Delta_1} = 0 \end{aligned} \quad (15)$$

и условиям на бесконечности. ■

Решение задачи (14), (15) в областях D_1, D_2 представим через базисные поля (6)

$$\begin{aligned} \vec{E}'_1(\lambda) &= x_1(\lambda)\vec{M}_0^{(+1)} + y_1(\lambda)\vec{M}_0^{(+2)}, \quad \vec{H}'_1(\lambda) = h_0(x_1(\lambda)\vec{M}_0^{(+2)} + y_1(\lambda)\vec{M}_0^{(+1)}), \\ \vec{E}_2(\lambda) &= x_2(\lambda)\vec{M}_0^{(-1)} + y_2(\lambda)\vec{M}_0^{(-2)}, \quad \vec{H}_2(\lambda) = h_0(x_2(\lambda)\vec{M}_0^{(-2)} + y_2(\lambda)\vec{M}_0^{(-1)}). \end{aligned} \quad (16)$$

Поля в слоях Ω_s исключим из рассмотрения, заменяя граничные условия (15) на двухсторонние граничные условия, связывающие поля $\vec{E}_1, \vec{H}_1; \vec{E}_2, \vec{H}_2$ по обе стороны экрана D . Для формулировки двухсторонних граничных условий запишем тангенциальные составляющие полей (13), (16) на плоскостях Γ_1, Γ_2 в базисе \vec{V}_1, \vec{V}_2 (7):

$$\begin{aligned} \vec{E}_{1\tau}(\lambda)\Big|_{z=0} &= E_{1V_1}(\lambda)\vec{V}_1 + E_{1V_2}(\lambda)\vec{V}_2, \quad \vec{H}_{1\tau}(\lambda)\Big|_{z=0} = H_{1V_1}(\lambda)\vec{V}_1 + H_{1V_2}(\lambda)\vec{V}_2, \\ \vec{E}_{2\tau}(\lambda)\Big|_{z=\Delta} &= E_{2V_1}(\lambda)\vec{V}_1 + E_{2V_2}(\lambda)\vec{V}_2, \quad \vec{H}_{2\tau}(\lambda)\Big|_{z=\Delta} = H_{2V_1}(\lambda)\vec{V}_1 + H_{2V_2}(\lambda)\vec{V}_2, \end{aligned} \quad (17)$$

где

$$\begin{aligned} E_{1V_1} &= E'_{1V_1} + E_{0V_1} = x_1(\lambda) + a(\lambda), \quad E_{1V_2} = E'_{1V_2} + E_{0V_2} = \bar{v}_0(\lambda)(y_1(\lambda) - b(\lambda)), \\ H_{1V_1} &= H'_{1V_1} + H_{0V_1} = h_0(y_1(\lambda) + b(\lambda)), \quad H_{1V_2} = H'_{1V_2} + H_{0V_2} = h_0\bar{v}_0(\lambda)(x_1(\lambda) - a(\lambda)), \end{aligned} \quad (18)$$

$$E_{2V_1} = x_2(\lambda)F(\lambda), \quad E_{2V_2} = -\bar{v}_0(\lambda)y_2(\lambda)F(\lambda), \quad \bar{v}_0(\lambda) = \frac{v_0(\lambda)}{k_0},$$

$$H_{2V_1} = h_0y_2(\lambda)F(\lambda), \quad H_{2V_2} = -h_0\bar{v}_0(\lambda)x_2(\lambda)F(\lambda), \quad F(\lambda) = e^{-v_0(\lambda)\Delta}.$$

Сформулируем краевую задачу проникновения поля (13) через экран D (с использованием двухсторонних граничных условий), которая эквивалентна краевой задаче (14), (15).

Краевая задача 3. При заданном поле \vec{E}_0, \vec{H}_0 (13) требуется определить поля $\vec{E}'_1, \vec{H}'_1; \vec{E}_2, \vec{H}_2$, которые удовлетворяют уравнениям

$$\operatorname{rot} \vec{E}_j = i\omega\mu_0 \vec{H}_j, \operatorname{rot} \vec{H}_j = -i\omega\varepsilon_0 \vec{E}_j \quad \text{в } D_j; \quad (19)$$

двухсторонним граничным условиям

$$\vec{U}_1|_{z=0} = \hat{S}(\lambda)\vec{U}_2|_{z=\Delta} \quad (20)$$

и условиям излучения в областях D_j .

Матрица $\hat{S}(\lambda)$ определяется как произведение двух матриц:

$$\hat{S}(\lambda) = \hat{C}_1(\lambda)\hat{C}_2(\lambda) = \begin{pmatrix} S_{11} & S_{12} & S_{13} & S_{14} \\ S_{21} & S_{22} & S_{23} & S_{24} \\ S_{31} & S_{32} & S_{33} & S_{34} \\ S_{41} & S_{42} & S_{43} & S_{44} \end{pmatrix}, \vec{U}_j = \begin{pmatrix} E_{jV_1} \\ H_{jV_2} \\ E_{jV_2} \\ H_{jV_1} \end{pmatrix}, \quad (21)$$

где E_{jv_s}, H_{jv_s} — компоненты полей (17).

4. Вычисление передаточной матрицы для двухслойного экрана

Элементы передаточной матрицы

$$\hat{C}_s = \begin{pmatrix} C_{11}^{(s)} & C_{12}^{(s)} & C_{13}^{(s)} & C_{14}^{(s)} \\ C_{21}^{(s)} & C_{22}^{(s)} & C_{23}^{(s)} & C_{24}^{(s)} \\ C_{31}^{(s)} & C_{32}^{(s)} & C_{33}^{(s)} & C_{34}^{(s)} \\ C_{41}^{(s)} & C_{42}^{(s)} & C_{43}^{(s)} & C_{44}^{(s)} \end{pmatrix}, \quad s = 1, 2, \quad (22)$$

для слоя Ω_s определяются формулами [7]

$$\begin{aligned} C_{11}^{(s)} &= p_s(p_{1s}C_2^{(s)} - p_{2s}C_1^{(s)}), \quad C_{12}^{(s)} = p_s\left(q_{1s}\frac{k_{1s}}{v_{1s}}S_1^{(s)} - q_{2s}\frac{k_{2s}}{v_{2s}}S_2^{(s)}\right), \\ C_{13}^{(s)} &= p_s\left(q_{2s}\frac{p_{1s}k_{2s}}{v_{2s}}S_2^{(s)} - q_{1s}\frac{p_{2s}k_{1s}}{v_{1s}}S_1^{(s)}\right), \quad C_{14}^{(s)} = p_s(C_1^{(s)} - C_2^{(s)}), \\ C_{21}^{(s)} &= p_s p_{1s} p_{2s} \left(q_{2s} \frac{v_{2s}}{k_{2s}} S_2^{(s)} - q_{1s} \frac{v_{1s}}{k_{1s}} S_1^{(s)} \right), \quad C_{22}^{(s)} = p_s(p_{1s}C_1^{(s)} - p_{2s}C_2^{(s)}), \\ C_{23}^{(s)} &= p_s p_{1s} p_{2s} (C_2^{(s)} - C_1^{(s)}), \quad C_{24}^{(s)} = p_s \left(\frac{p_{1s}v_{1s}}{q_{1s}k_{1s}} S_1^{(s)} - \frac{p_{2s}v_{2s}}{q_{2s}k_{2s}} S_2^{(s)} \right), \\ C_{31}^{(s)} &= p_s \left(\frac{p_{1s}v_{2s}}{q_{2s}k_{2s}} S_2^{(s)} - \frac{p_{2s}v_{1s}}{q_{1s}k_{1s}} S_1^{(s)} \right), \quad C_{32}^{(s)} = p_s(C_1^{(s)} - C_2^{(s)}), \\ C_{33}^{(s)} &= p_s(p_{1s}C_2^{(s)} - p_{2s}C_1^{(s)}), \quad C_{34}^{(s)} = p_s \left(\frac{v_{1s}}{q_{1s}k_{1s}} S_1^{(s)} - \frac{v_{2s}}{q_{2s}k_{2s}} S_2^{(s)} \right), \\ C_{41}^{(s)} &= p_s p_{1s} p_{2s} (C_2^{(s)} - C_1^{(s)}), \quad C_{42}^{(s)} = p_s \left(q_{1s} p_{1s} \frac{k_{1s}}{v_{1s}} S_1^{(s)} - q_{2s} p_{2s} \frac{k_{2s}}{v_{2s}} S_2^{(s)} \right), \end{aligned} \quad (23)$$

$$C_{43}^{(s)} = p_s p_{1s} p_{2s} \left(q_{2s} \frac{k_{2s}}{v_{2s}} S_2^{(s)} - q_{1s} \frac{k_{1s}}{v_{1s}} S_1^{(s)} \right), \quad C_{44}^{(s)} = p_s (p_{1s} C_1^{(s)} - p_{2s} C_2^{(s)}), \quad (23)$$

$$C_j^{(s)} = \text{ch}(\xi_{js}), \quad S_j^{(s)} = \text{sh}(\xi_{js}), \quad \xi_{js} = v_{js}(\lambda) \Delta_s, \quad v_{js}(\lambda) = \sqrt{\lambda^2 - k_{js}^2}, \quad -\frac{\pi}{2} \leq \arg v_j < \frac{\pi}{2}.$$

Величины в формулах (23)

$$k_{js} = k_0 \bar{k}_{js}, \quad f_{0s} = k_0 \bar{f}_{0s}, \quad f_{js} = k_0 \bar{f}_{js}, \quad g_s = k_0^2 \bar{g}_s, \quad g_j = k_0 \bar{g}_j, \quad a_s = k_0 \bar{a}_s, \quad b_s = k_0 \bar{b}_s, \quad p_{js} = \bar{p}_{js} / Z_0,$$

$$q_{js} = \frac{\bar{g}_s}{k_{js} \bar{g}_{js}}, \quad v_{js} = k_0 \bar{v}_{js}, \quad p_s = \frac{1}{p_{1s} - p_{2s}} = Z_0 \bar{p}_s, \quad k_0 = \frac{\omega}{c}$$

выразим через безразмерные величины, обозначаемые чертой сверху:

$$\begin{aligned} \bar{k}_{js} &= \sqrt{\bar{g}_s + \frac{1}{2} \bar{a}_s^2 + \bar{a}_s \bar{f}_{js}}, \quad \bar{f}_{js} = (-1)^j \bar{f}_{0s}, \quad \bar{f}_{0s} = \sqrt{\varepsilon_{rs} \mu_{rs} - \bar{b}_s^2}, \\ 0 \leq \arg \bar{k}_{js} &< \pi, \quad 0 \leq \arg \bar{f}_{0s} < \pi, \quad \bar{g}_s = \varepsilon_{rs} \mu_{rs} - G_{rs} Z_{rs}, \end{aligned} \quad (24)$$

$$\bar{g}_{js} = \bar{f}_{js} - \frac{1}{2} \bar{a}_s, \quad \bar{p}_{js} = \frac{1}{\mu_{rs}} \left(\frac{i \bar{g}_s}{\bar{g}_{js}} - Z_{rs} \right), \quad \bar{a}_s = i(G_{rs} - Z_{rs}),$$

$$\bar{b}_s = \frac{1}{2} (G_{rs} + Z_{rs}), \quad \bar{v}_{js}(\lambda) = \sqrt{\bar{\lambda} - \bar{k}_{js}^2}, \quad -\frac{\pi}{2} \leq \arg \bar{v}_{js} < \frac{\pi}{2}, \quad \bar{\lambda} = \frac{\lambda}{k_0},$$

$$\bar{p}_s = \frac{1}{\bar{p}_{1s} - \bar{p}_{2s}}.$$

Наложим на материальные параметры слоев Ω_s условия фокусировки

$$\varepsilon_{rs} \mu_{rs} - \chi_s^2 = 1, \quad G_{rs} = Z_{rs} = \chi_s, \quad s = 1, 2, \quad (25)$$

где χ_s – произвольные комплексные числа.

Тогда безразмерные величины (24) примут значения

$$\bar{a}_s = 0, \quad \bar{b}_s = \chi_s, \quad \bar{g}_s = 1, \quad \bar{k}_{1s} = \bar{k}_{2s} = 1, \quad \bar{g}_{1s} = -1, \quad \bar{g}_{2s} = 1, \quad f_{0s} = 1, \quad \bar{f}_{1s} = -1, \quad \bar{f}_{2s} = 1,$$

$$q_{1s} = -1, \quad q_{2s} = 1, \quad \bar{p}_{1s} = -\frac{i}{\mu_{rs}} (1 - i\chi_s), \quad \bar{p}_{2s} = \frac{i}{\mu_s} (1 + i\chi_s), \quad \bar{p}_s = \frac{i\mu_s}{2}, \quad \bar{v}_{1s} = \bar{v}_{2s} = \bar{v}_0(\lambda).$$

Матричные элементы матрицы (22) упрощаются. Получим

$$\hat{C}_s = \begin{pmatrix} C_s & C_{12}^{(s)} & C_{13}^{(s)} & 0 \\ C_{21}^{(s)} & C_s & 0 & C_{24}^{(s)} \\ C_{31}^{(s)} & 0 & C_s & C_{34}^{(s)} \\ 0 & C_{42}^{(s)} & C_{43}^{(s)} & C_s \end{pmatrix}, \quad s = 1, 2, \quad (26)$$

где $C_{11}^{(s)} = C_{22}^{(s)} = C_{33}^{(s)} = C_{44}^{(s)} = C_s, \quad C_{14}^{(s)} = C_{23}^{(s)} = C_{32}^{(s)} = C_{41}^{(s)} = 0,$

$$\begin{aligned}
C_{12}^{(s)} &= -\frac{i\mu_{rs}}{\bar{v}_0} Z_0 S_s, & C_{13}^{(s)} &= -\frac{i\chi_s}{\bar{v}_0} S_s, & C_s &= \text{ch}(\xi_s), & S_s &= \text{sh}(\xi_s), & \xi_s &= v_0(\lambda)\Delta_s, \\
C_{21}^{(s)} &= \frac{i\varepsilon_{rs}}{Z_0} \bar{v}_0 S_s, & C_{24}^{(s)} &= i\chi_s \bar{v}_0 S_s, & C_{31}^{(s)} &= -i\chi_s \bar{v}_0 S_s, \\
C_{34}^{(s)} &= -iZ_0 \mu_{rs} \bar{v}_0 S_s, & C_{42}^{(s)} &= \frac{i\chi_s}{\bar{v}_0} S_s, & C_{43}^{(s)} &= \frac{i\varepsilon_{rs}}{Z_0 \bar{v}_0} S_s.
\end{aligned}$$

Вычислим матрицу (21), умножая матрицы (26). Для матричных элементов получим формулы

$$\begin{aligned}
S_{11} &= S_{33} = C_1 C_2 + (\mu_{r1} \varepsilon_{r2} - \chi_1 \chi_2) S_1 S_2, & S_{22} &= S_{44} = C_1 C_2 + (\varepsilon_{r1} \mu_{r2} - \chi_1 \chi_2) S_1 S_2, \\
S_{12} &= -\frac{iZ_0}{\bar{v}_0} (\mu_{r1} C_2 S_1 + \mu_{r2} C_1 S_2), & S_{13} &= -\frac{i}{\bar{v}_0} (\chi_1 C_2 S_1 + \chi_2 C_1 S_2), \\
S_{14} &= (\mu_{r1} \chi_2 - \mu_{r2} \chi_1) Z_0 S_1 S_2, & S_{21} &= \frac{i\bar{v}_0}{Z_0} (\varepsilon_{r1} C_2 S_1 + \varepsilon_{r2} C_1 S_2), \\
S_{23} &= \frac{1}{Z_0} (\varepsilon_{r1} \chi_2 - \varepsilon_{r2} \chi_1) S_1 S_2, & S_{24} &= i\bar{v}_0 (\chi_1 C_2 S_1 + \chi_2 C_1 S_2), \\
S_{34} &= -i\bar{v}_0 Z_0 (\mu_{r1} C_2 S_1 + \mu_{r2} C_1 S_2), & S_{43} &= \frac{i}{Z_0 \bar{v}_0} (\varepsilon_{r1} C_2 S_1 + \varepsilon_{r2} C_1 S_2), \\
S_{31} &= -S_{24}, & S_{32} &= S_{14}, & S_{41} &= S_{23}, & S_{42} &= -S_{13}.
\end{aligned} \tag{27}$$

Матрица (21) граничного условия определена.

5. Аналитическое решение краевых задач экранирования

Решение краевой задачи 3 представим в виде полей (16), которые удовлетворяют уравнениям (19). Для определения неизвестных функций $x_s(\lambda)$, $y_s(\lambda)$ запишем векторное граничное условие (20) в виде четырех скалярных уравнений, подставляя в (20) компоненты (18) векторов \vec{U}_s . Получим систему алгебраических уравнений для вычисления x_s , y_s :

$$\begin{aligned}
(S_{11} - h_0 \bar{v}_0 S_{12}) x_2 + (h_0 S_{14} - \bar{v}_0 S_{13}) y_2 &= (x_1 + a)/F, \\
(S_{21} - h_0 \bar{v}_0 S_{22}) x_2 + (h_0 S_{24} - \bar{v}_0 S_{23}) y_2 &= h_0 \bar{v}_0 (x_1 - a)/F, \\
(S_{31} - h_0 \bar{v}_0 S_{32}) x_2 + (h_0 S_{34} - \bar{v}_0 S_{33}) y_2 &= \bar{v}_0 (y_1 - b)/F, \\
(S_{41} - h_0 \bar{v}_0 S_{42}) x_2 + (h_0 S_{44} - \bar{v}_0 S_{43}) y_2 &= h_0 (y_1 + b)/F.
\end{aligned} \tag{28}$$

Решая систему уравнений (28), получим

$$\begin{aligned}
x_2 &= \frac{2}{DF} (aK_{22} - bK_{12}), & y_2 &= \frac{2}{DF} (bK_{11} - aK_{21}), \\
x_1 &= \frac{2}{D} [(S_{11} - h_0 \bar{v}_0 S_{12})(aK_{22} - bK_{12}) + (\bar{v}_0 S_{13} - h_0 S_{14})(aK_{21} - bK_{11})] - a, \\
y_1 &= \frac{2}{D} [(iZ_0 S_{41} - \bar{v}_0 S_{42})(aK_{22} - bK_{12}) + (iZ_0 \bar{v}_0 S_{43} - S_{44})(aK_{21} - bK_{11})] - b,
\end{aligned} \tag{29}$$

где с учетом (27)

$$K_{11} = S_{11} - h_0 \bar{v}_0 S_{12} - \frac{iZ_0}{\bar{v}_0} S_{21} + S_{22} = K(\lambda) = \beta_0 S_1 S_2 + \alpha_1 C_2 S_1 + \alpha_2 C_1 S_2 + 2C_1 C_2,$$

$$K_{22} = S_{33} - \frac{h_0}{\bar{v}_0} S_{34} - iZ_0 \bar{v}_0 S_{43} + S_{44} = K(\lambda),$$

$$K_{12} = -\bar{v}_0 S_{13} + h_0 S_{14} + iZ_0 S_{23} - \frac{1}{\bar{v}_0} S_{24} = i\alpha_0 S_1 S_2 = K_0(\lambda),$$

$$K_{21} = -\frac{1}{\bar{v}_0} S_{31} + h_0 S_{32} + iZ_0 S_{41} - \bar{v}_0 S_{42} = K_0(\lambda),$$

$$D = K_{11} K_{22} - K_{12} K_{21} = K^2 - K_0^2, \quad \alpha_0 = (\varepsilon_{r1} - \mu_{r1}) \chi_2 + (\mu_{r2} - \varepsilon_{r2}) \chi_1,$$

$$\beta_0 = \mu_{r1} \varepsilon_{r2} + \varepsilon_{r1} \mu_{r2} - 2\chi_1 \chi_2, \quad \alpha_1 = \varepsilon_{r1} + \mu_{r1}, \quad \alpha_2 = \varepsilon_{r2} + \mu_{r2}.$$

Решение (16) краевой задачи (19)–(21) при условиях фокусировки (25) построено.

Для магнитного диполя (13) $b = 0$. Подставляя функции (29) в (16) и далее в интегралы (12), определим электромагнитные поля $\vec{E}'_1, \vec{H}'_1; \vec{E}_2, \vec{H}_2$, являющиеся решением краевой задачи 1. Для электрических полей получим следующие интегральные представления:

$$\vec{E}_2 = \int_0^\infty \frac{2A(\lambda, -h)}{D(\lambda)F(\lambda)} K(\lambda) \vec{M}_0^{(-1)}(\vec{\rho}; \lambda, k_0) d\lambda - \int_0^\infty \frac{2A(\lambda, -h)}{D(\lambda)F(\lambda)} K_0(\lambda) \vec{M}_0^{(-2)}(\vec{\rho}; \lambda, k_0) d\lambda, \quad (30)$$

$$\vec{E}'_1 = \int_0^\infty 2 \frac{A(\lambda, -h)}{D(\lambda)} P(\lambda) \vec{M}_0^{(+1)}(\vec{\rho}; \lambda, k_0) d\lambda + \int_0^\infty 2 \frac{A(\lambda, -h)}{D(\lambda)} P_0(\lambda) \vec{M}_0^{(+2)}(\vec{\rho}; \lambda, k_0) d\lambda,$$

где

$$P(\lambda) = (S_{11} - h_0 \bar{v}_0 S_{12})K + (\bar{v}_0 S_{13} - h_0 S_{14})K_0 - \frac{1}{2}D,$$

$$P_0(\lambda) = (iZ_0 S_{41} - \bar{v}_0 S_{42})K + (iZ_0 \bar{v}_0 S_{43} - S_{44})K_0.$$

6. Преобразование первичного TE -поля в TE -поле за экраном

Первичное поле (13) является TE -полем, которое при прохождении за экран в область D_2 преобразуется в поле (30). Поле (30) определяется суммой TE -поля и TH -поля. Исключим TH -поле, накладывая на материальные параметры экрана дополнительное условие фокусировки

$$\alpha_0 = (\varepsilon_{r1} - \mu_{r1}) \chi_2 + (\mu_{r2} - \varepsilon_{r2}) \chi_1 = 0, \quad (31)$$

тогда

$$K_0 = 0, \quad D = K^2, \quad P = \bar{P}K, \quad \bar{P} = S_{11} - h_0 \bar{v}_0 S_{12} - \frac{1}{2}K, \quad P_0 = \bar{P}_0 K, \quad \bar{P}_0 = iZ_0 S_{41} - \bar{v}_0 S_{42}.$$

Запишем эти величины через экспоненты:

$$K = \frac{1}{4} \left(\tau_1 e^{-v_0 \Delta} + \tau_2 e^{v_0(\Delta_2 - \Delta_1)} + \tau_3 e^{v_0(\Delta_1 - \Delta_2)} + \tau_4 e^{v_0 \Delta} \right), \quad (32)$$

$$\bar{P}_0 = \frac{i}{4} \left(\nu_1 e^{-v_0 \Delta} + \nu_2 e^{v_0(\Delta_2 - \Delta_1)} + \nu_3 e^{v_0(\Delta_1 - \Delta_2)} + \nu_4 e^{v_0 \Delta} \right),$$

$$\bar{P} = \frac{1}{8} \left(\sigma_1 e^{-\nu_0 \Delta} + \sigma_2 e^{\nu_0 (\Delta_2 - \Delta_1)} + \sigma_3 e^{\nu_0 (\Delta_1 - \Delta_2)} + \sigma_4 e^{\nu_0 \Delta} \right),$$

где

$$\tau_1 = 2 + \beta_0 - \alpha_1 - \alpha_2, \quad \tau_2 = 2 - \beta_0 - \alpha_1 + \alpha_2,$$

$$\tau_3 = 2 - \beta_0 + \alpha_1 - \alpha_2, \quad \tau_4 = 2 + \beta_0 + \alpha_1 + \alpha_2;$$

$$\nu_1 = (\varepsilon_{r1} + 1)\chi_2 - (\varepsilon_{r2} - 1)\chi_1, \quad \nu_2 = (\varepsilon_{r2} + 1)\chi_1 - (\varepsilon_{r1} + 1)\chi_2,$$

$$\nu_3 = (\varepsilon_{r2} - 1)\chi_1 - (\varepsilon_{r1} - 1)\chi_2, \quad \nu_4 = (\varepsilon_{r1} - 1)\chi_2 - (\varepsilon_{r2} + 1)\chi_1;$$

$$\sigma_1 = -\gamma_0 + \alpha_1 + \alpha_2, \quad \sigma_2 = \gamma_0 + \alpha_1 - \alpha_2, \quad \sigma_3 = \gamma_0 - \alpha_1 + \alpha_2, \quad \sigma_4 = -\gamma_0 - \alpha_1 - \alpha_2,$$

$$\gamma_0 = \varepsilon_{r1} \mu_{r2} - \mu_{r1} \varepsilon_{r2}.$$

В результате при условии (31) поля (30) примут вид

$$\bar{\mathbf{E}}_2 = 2 \int_0^\infty \frac{A(\lambda, -h)}{K(\lambda) F(\lambda)} \bar{M}_0^{(-1)}(\bar{\rho}; \lambda, k_0) d\lambda; \quad (33)$$

$$\bar{\mathbf{E}}_1' = 2 \int_0^\infty \frac{A(\lambda, -h)}{K(\lambda)} \bar{P}(\lambda) \bar{M}_0^{(+1)}(\bar{\rho}; \lambda, k_0) d\lambda + 2 \int_0^\infty \frac{A(\lambda, -h)}{K(\lambda)} \bar{P}_0(\lambda) \bar{M}_0^{(+2)}(\bar{\rho}; \lambda, k_0) d\lambda. \quad (34)$$

7. Способ фокусировки полей 1

Для фокусировки электромагнитных полей экраном D наложим на материальные параметры слоев экрана дополнительные условия фокусировки. В выражении (32) при условии $\Delta_1 \neq \Delta_2$ положим

$$\tau_2 = 2 - \beta_0 - \alpha_1 + \alpha_2 = 0, \quad \tau_3 = 2 - \beta_0 + \alpha_1 - \alpha_2 = 0, \quad \tau_4 = 2 + \beta_0 + \alpha_1 + \alpha_2 = 0. \quad (35)$$

Разрешим систему алгебраических уравнений (35):

$$\alpha_1 = \varepsilon_{r1} + \mu_{r1} = -2, \quad \alpha_2 = \varepsilon_{r2} + \mu_{r2} = -2, \quad \beta_0 = \varepsilon_{r1} \mu_{r2} + \mu_{r1} \varepsilon_{r2} - 2\chi_1 \chi_2 = 2. \quad (36)$$

В результате получим систему уравнений фокусировки (25), (31), (36), которая имеет два линейных многообразия решений:

– первое многообразие

$$\varepsilon_{r1} = \delta_1, \quad \mu_{r1} = -2 - \delta_1, \quad G_{r1} = Z_{r1} = i(1 + \delta_1), \quad (37)$$

$$\varepsilon_{r2} = \delta_2, \quad \mu_{r2} = -2 - \delta_2, \quad G_{r2} = Z_{r2} = i(1 + \delta_2),$$

где δ_1, δ_2 – произвольные комплексные величины;

– второе многообразие

$$\varepsilon_{r1} = \delta_1, \quad \mu_{r1} = -2 - \delta_1, \quad G_{r1} = Z_{r1} = -i(1 + \delta_1), \quad (38)$$

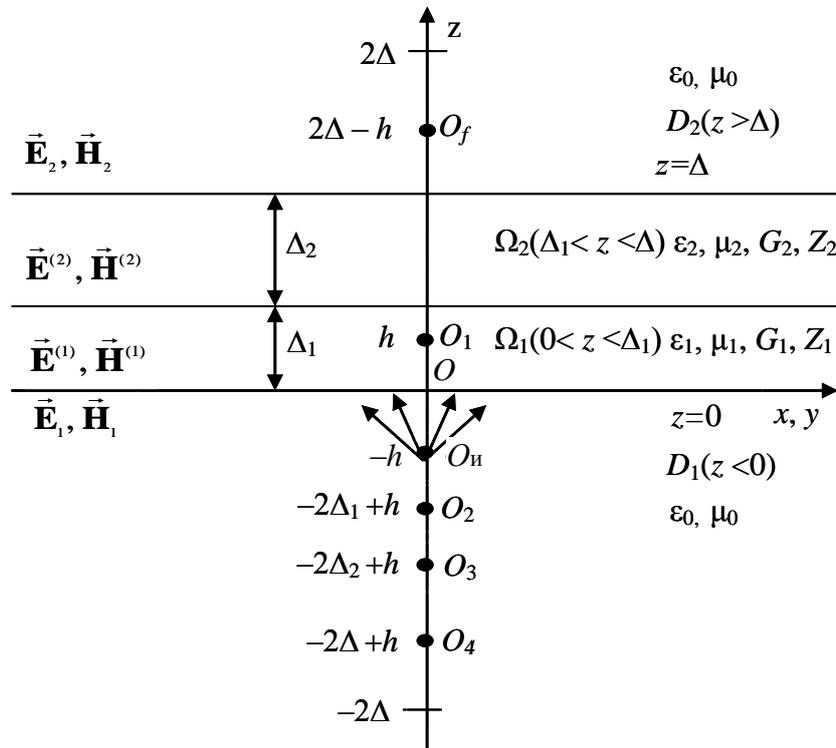
$$\varepsilon_{r2} = \delta_2, \quad \mu_{r2} = -2 - \delta_2, \quad G_{r2} = Z_{r2} = -i(1 + \delta_2).$$

Таким образом, если экран выполнен из материалов (37) или (38), то $K(\lambda) = \frac{\tau_1}{4} e^{-\nu_0(\lambda)\Delta}$ и поле за экраном (33) с учетом формул (9), (10) определяется выражением

$$\begin{aligned} \vec{\mathbf{E}}_2 &= \frac{8}{\tau_1} \int_0^\infty A(\lambda, -h) e^{2\nu_0(\lambda)\Delta} \vec{M}_0^{(-1)}(\vec{\rho}; \lambda, k_0) d\lambda = \\ &= \frac{8}{\tau_1} \int_0^\infty A(\lambda, h_f) \vec{M}_0^{(-1)}(\vec{\rho}; \lambda, k_0) d\lambda = \frac{8}{\tau_1} \vec{\mathbf{E}}_m(O_f) = A_0 \tilde{m}_{01}(\vec{r}_f, k_0), \end{aligned} \quad (39)$$

где $h_f = 2\Delta - h$, с учетом (32), (36) $\tau_1 = 8$.

Формула (39) показывает, что поле в области D_2 является полем магнитного диполя, сосредоточенного в точке O_f с координатами $x=0, y=0, z=h_f$. Из условия $0 < h < \Delta$ следует, что $h_f > \Delta$, т. е. источник O_f расположен в области D_2 на расстоянии $\Delta - h$ от экрана D и является фокусом (рисунок).



Двухслойная линза с точками фокусировки. Способ фокусировки 1, $0 < h < 2\Delta_1, 0 < h < 2\Delta_2, 0 < h < \Delta$; $O_{и}$ – источник поля, O_f, O_2, O_3, O_4 – фокусы, O_1 – фиктивный фокус

Вычислим отраженное поле (34) в области D_1 :

$$\begin{aligned} \vec{\mathbf{E}}'_1 &= \frac{1}{\tau_1} \int_0^\infty A(\lambda, -h) e^{\nu_0\Delta} \left(\sigma_1 e^{-\nu_0\Delta} + \sigma_2 e^{\nu_0(\Delta_2-\Delta_1)} + \sigma_3 e^{\nu_0(\Delta_1-\Delta_2)} + \sigma_4 e^{\nu_0\Delta} \right) \vec{M}_0^{(+1)}(\vec{\rho}; \lambda, k_0) d\lambda + \\ &+ \frac{2i}{\tau_1} \int_0^\infty A(\lambda, -h) e^{\nu_0\Delta} \left(\nu_1 e^{-\nu_0\Delta} + \nu_2 e^{\nu_0(\Delta_2-\Delta_1)} + \nu_3 e^{\nu_0(\Delta_1-\Delta_2)} + \nu_4 e^{-\nu_0\Delta} \right) \vec{M}_0^{(+2)}(\vec{\rho}; \lambda, k_0) d\lambda. \end{aligned}$$

Учитывая формулы (4), (5), (9), (10), получим представление поля

$$\begin{aligned} \vec{\mathbf{E}}'_1 &= \frac{1}{8} \int_0^\infty \left(\sum_{l=1}^4 \sigma_l A(\lambda, -h_l) \right) \vec{M}_0^{(+1)}(\vec{\rho}; \lambda, k_0) d\lambda + \frac{i}{4} \int_0^\infty \left(\sum_{l=1}^4 \nu_l A(\lambda, -h_l) \right) \vec{M}_0^{(+2)}(\vec{\rho}; \lambda, k_0) d\lambda = \\ &= \sum_{l=1}^4 \left(\frac{1}{8} \sigma_l \vec{\mathbf{E}}_m(O_l) + \frac{i}{4} \nu_l \vec{\mathbf{E}}_s(O_l) \right) = A_0 \sum_{l=1}^4 \left(\frac{1}{8} \sigma_l \tilde{m}_{01}(O_l) + \frac{i}{4} \nu_l \tilde{n}_{01}(O_l) \right), \end{aligned} \quad (40)$$

где

$$\begin{aligned} h_1 &= h, h_2 = h - 2\Delta_2, h_3 = h - 2\Delta_1, h_4 = h - 2\Delta; \\ \sigma_1 &= 2(\delta_1 - \delta_2) - 4, \sigma_2 = \sigma_3 = 2(\delta_2 - \delta_1), \sigma_4 = 2(\delta_1 - \delta_2) + 4; \\ v_1 &= 2i(1 + \delta_1), v_2 = 0, v_3 = 2i(\delta_2 - \delta_1), v_4 = -2i(1 + \delta_2). \end{aligned}$$

С учетом формул (37) и (38) $v_1 = -2i(1 + \delta_1)$, $v_2 = 0$, $v_3 = -2i(\delta_2 - \delta_1)$, $v_4 = 2i(1 + \delta_2)$.

Анализ формул (40) показывает, что отраженное поле в области D_1 состоит из суммы полей электрических и магнитных диполей, сосредоточенных в четырех точках O_l ($l=1, 2, 3, 4$). Точка O_1 во всех случаях расположена вне полупространства D_1 , а при $h_l < 0$ источник O_l поля диполя принадлежит области D_1 и является фокусом, который формируется перед экраном в области источника O_{II} первичного поля. В частности, при выполнении условий $0 < h < 2\Delta_1$, $0 < h < 2\Delta_2$ точки O_2 , O_3 , O_4 являются фокусами в области D_1 (см. рисунок); при $0 < h < 2\Delta_1$, $h > 2\Delta_2$ точки O_3 , O_4 – фокусы в области D_1 ; при $0 < h < 2\Delta_2$, $h > 2\Delta_1$ точки O_2 , O_4 – фокусы в области D_1 ; при $2\Delta_1 < h$, $2\Delta_2 < h$, $h < 2\Delta$ точка O_4 – фокус.

8. Способ фокусировки полей 2

Вместо условий (35) выберем другую комбинацию условий фокусировки. В выражении (32) положим

$$\tau_1 = 2 + \beta_0 - \alpha_1 - \alpha_2 = 0, \tau_3 = 2 - \beta_0 + \alpha_1 - \alpha_2 = 0, \tau_4 = 2 + \beta_0 + \alpha_1 + \alpha_2 = 0.$$

Решая эту систему уравнений, получим

$$\alpha_1 = \varepsilon_{r1} + \mu_{r1} = -2, \alpha_2 = \varepsilon_{r2} + \mu_{r2} = 2, \beta_0 = \varepsilon_{r1}\mu_{r2} + \mu_{r1}\varepsilon_{r2} - 2\chi_1\chi_2 = -2. \quad (41)$$

Система уравнений фокусировки (25), (31), (41) имеет два линейных многообразия решений:

$$\varepsilon_{r1} = \delta_1, \quad \mu_{r1} = -2 - \delta_1, \quad G_{r1} = Z_{r1} = i(\delta_1 + 1), \quad (42)$$

$$\varepsilon_{r2} = \delta_2, \quad \mu_{r2} = 2 - \delta_2, \quad G_{r2} = Z_{r2} = i(\delta_2 - 1);$$

$$\varepsilon_{r1} = \delta_1, \quad \mu_{r1} = -2 - \delta_1, \quad G_{r1} = Z_{r1} = -i(\delta_1 + 1), \quad (43)$$

$$\varepsilon_{r2} = \delta_2, \quad \mu_{r2} = 2 - \delta_2, \quad G_{r2} = Z_{r2} = -i(\delta_2 - 1),$$

где δ_1, δ_2 – произвольные комплексные величины.

Если экран выполнен из материалов (42) или (43), то $K(\lambda) = \frac{\tau_2}{4} e^{v_0(\lambda)(\Delta_2 - \Delta_1)}$ и поле за экраном (33) с учетом формул (9), (10) определяется выражением

$$\begin{aligned} \vec{E}_2 &= \frac{8}{\tau_2} \int_0^\infty A(\lambda, -h) e^{v_0(\lambda)(\Delta + \Delta_1 - \Delta_2)} \vec{M}_0^{(-1)}(\vec{\rho}; \lambda, k_0) d\lambda = \\ &= \frac{8}{\tau_2} \int_0^\infty A(\lambda, h_{f1}) \vec{M}_0^{(-1)}(\vec{\rho}; \lambda, k_0) \vec{M}_0^{(-1)}(\vec{\rho}; \lambda, k_0) d\lambda = \frac{8}{\tau_2} \vec{E}_m(O_{f1}) = A_0 \vec{m}_{01}(\vec{r}_{f1}, k_0), \end{aligned}$$

где $h_{f1} = \Delta + \Delta_1 - \Delta_2 - h$, $\tau_2 = 8$.

Показано, что поле $\vec{\mathbf{E}}_2$ в области D_2 является полем магнитного диполя, сосредоточенного в точке $O_{f_1}(0,0,h_{f_1})$. Из условия $0 < h < \Delta_1 - \Delta_2$ ($\Delta_1 > \Delta_2$) следует $h_{f_1} > \Delta$, т. е. источник O_{f_1} расположен в области D_2 на расстоянии $\Delta_1 - \Delta_2 - h$ от экрана и является фокусом.

Отраженное поле в области D_1 определяется формулами (40), где для источников поля $O_l(0,0,h_l)$ $h_1 = 2\Delta_2 + h$, $h_2 = h$, $h_3 = 2(\Delta_2 - \Delta_1) + h$, $h_4 = -2\Delta_1 + h$. Для материальных параметров (42) коэффициенты $\sigma_1 = \sigma_4 = -2(\delta_1 + \delta_2)$, $\sigma_2 = 2(\delta_1 + \delta_2) - 4$, $\sigma_3 = 2(\delta_1 + \delta_2) + 4$; $v_1 = 0$, $v_2 = 2i(\delta_1 + 1)$, $v_3 = 2i(\delta_2 - 1)$, $v_4 = 2i(\delta_1 + \delta_2)$, а для параметров (43) $v_1 = 0$, $v_2 = -2i(\delta_1 + 1)$, $v_3 = -2i(\delta_2 - 1)$, $v_4 = -2i(\delta_1 + \delta_2)$.

Из значений координат точек O_l следует, что точки O_1 и O_2 во всех случаях лежат вне области D_1 . При условии $0 < h < 2(\Delta_1 - \Delta_2)$ точки O_3 и O_4 принадлежат области D_1 и являются фокусами.

При условии $2(\Delta_1 - \Delta_2) < h < 2\Delta_1$ точка O_3 лежит вне области D_1 , а точка $O_4 \in D_1$, т. е. является фокусом. При условии $h > 2\Delta_1$ точки O_3 и O_4 не принадлежат области D_1 .

9. Способ фокусировки полей 3

Выберем условия фокусировки, полагая в выражении (32)

$$\tau_1 = 2 + \beta_0 - \alpha_1 - \alpha_2 = 0, \quad \tau_2 = 2 - \beta_0 - \alpha_1 + \alpha_2 = 0, \quad \tau_4 = 2 + \beta_0 + \alpha_1 + \alpha_2 = 0.$$

Решив эту систему уравнений, получим

$$\alpha_1 = \varepsilon_{r1} + \mu_{r1} = 2, \quad \alpha_2 = \varepsilon_{r2} + \mu_{r2} = -2, \quad \beta_0 = \varepsilon_{r1}\mu_{r2} + \mu_{r1}\varepsilon_{r2} - 2\chi_1\chi_2 = -2. \quad (44)$$

Решив систему уравнений (25), (31), (44), получим два многообразия решений, которые определяют фокусирующие материальные параметры слоев:

$$\varepsilon_{r1} = \delta_1, \quad \mu_{r1} = 2 - \delta_1, \quad G_{r1} = Z_{r1} = i(\delta_1 - 1), \quad (45)$$

$$\varepsilon_{r2} = \delta_2, \quad \mu_{r2} = -2 - \delta_2, \quad G_{r2} = Z_{r2} = i(\delta_2 + 1);$$

$$\varepsilon_{r1} = \delta_1, \quad \mu_{r1} = 2 - \delta_1, \quad G_{r1} = Z_{r1} = -i(\delta_1 - 1), \quad (46)$$

$$\varepsilon_{r2} = \delta_2, \quad \mu_{r2} = -2 - \delta_2, \quad G_{r2} = Z_{r2} = -i(\delta_2 + 1),$$

где δ_1, δ_2 – произвольные комплексные величины.

Расположение фокусов определяется по аналогии со способом фокусировки 2. В результате

$$\vec{\mathbf{E}}_2 = \vec{\mathbf{E}}_m(O_{f_2}) = A_0 \tilde{m}_{01}(\vec{r}_{f_2}, k_0),$$

где координаты фокуса $O_{f_2} = (0,0,h_{f_2})$, $h_{f_2} = \Delta + \Delta_2 - \Delta_1 - h$;

$$\vec{\mathbf{E}}'_1 = \sum_{l=1}^4 \left(\frac{1}{8} \sigma_l \vec{\mathbf{E}}_m(O_l) + \frac{i}{4} v_l \vec{\mathbf{E}}_s(O_l) \right) = A_0 \sum_{l=1}^4 \left(\frac{1}{8} \sigma_l \tilde{m}_{01}(\vec{r}_l, k_0) + \frac{i}{4} v_l \tilde{n}_{01}(\vec{r}_l, k_0) \right).$$

Здесь для материальных параметров экрана (43) $\sigma_1 = \sigma_4 = 2(\delta_1 + \delta_2)$, $\sigma_2 = -2(\delta_1 + \delta_2) + 4$, $\sigma_3 = -2(\delta_1 + \delta_2) - 4$; $v_1 = 2i(\delta_1 + \delta_2)$, $v_2 = -2i(\delta_2 + 1)$, $v_3 = -2i(\delta_1 - 1)$, $v_4 = 0$; для материальных

параметров (44) коэффициенты σ_l совпадают с вышеприведенными коэффициентами, а $v_1 = -2i(\delta_1 + \delta_2)$, $v_2 = 2i(\delta_2 + 1)$, $v_3 = 2i(\delta_1 - 1)$, $v_4 = 0$.

Координаты точек расположения диполей $O_l = (0, 0, h_l)$ следующие:

$$h_1 = 2\Delta_1 + h, h_2 = 2(\Delta_1 - \Delta_2) + h, h_3 = h, h_4 = -2\Delta_2 + h.$$

Отметим, что дипольная структура фокусов по обе стороны однослойной линзы аналитически описана в материалах [17].

Заключение

Фокусирующие материальные параметры двух плоских слоев с толщинами Δ_1 и Δ_2 ($\Delta_1 + \Delta_2 = \Delta$, $\Delta_1 \neq \Delta_2$) двухслойной линзы (экрана) из метаматериалов составляют ряд одномерных линейных многообразий, параметризуемых независимыми произвольными комплексными величинами δ_1 и δ_2 . Возможны три основных варианта заполнения слоев фокусирующими материалами с фиксированными величинами δ_1 и δ_2 , при которых осуществляется фокусировка поля магнитного диполя, расположенного перед экраном на расстоянии h от экрана. Возможны другие варианты заполнения слоев. В случае типового варианта (способ фокусировки 1) имеем следующие фокусы: при $0 < h < \Delta$ за экраном образуется один фокус; при $0 < h < 2\Delta_1$, $0 < h < 2\Delta_2$ перед экраном возникают три фокуса; при условиях $0 < h < 2\Delta_1$, $2\Delta_2 < h$ или $0 < h < 2\Delta_2$, $2\Delta_1 < h$ – два фокуса; при $2\Delta_1 < h$, $2\Delta_2 < h$, $h < 2\Delta$ – один фокус; при $h > 2\Delta$ фокусы перед экраном отсутствуют. Рассмотрены и другие способы фокусировки.

Список литературы

1. Иванов, О.В. Распространение электромагнитных волн в анизотропных и бианизотропных слоистых структурах / О.И. Иванов. – Ульяновск : УлГТУ, 2010. – 262 с.
2. Неганов, В.А. Электродинамика отражающих и волноведущих структур с искусственными киральными слоями / В.А. Неганов, О.В. Осипов // Успехи современной радиоэлектроники. – 2005. – № 8. – С. 20–45.
3. Блиох, К.Ю. Что такое левые среды и чем они интересны / К.Ю. Блиох, Ю.П. Блиох // Успехи физических наук. – 2004. – Т. 174, № 4. – С. 439–447.
4. Electromagnetic scattering by multilayered chiral-media structures: a scattering – to – radiation transform / L.-W. Li [et al.] // Progress in electromagnetics research. – 2000. – PIER 26. – P. 249–291.
5. Ерофеенко, В.Т. Аналитическое моделирование в электродинамике / В.Т. Ерофеенко, И.С. Козловская. – Минск : БГУ, 2010. – 304 с.
6. Ерофеенко, В.Т. Дифракция плоской электромагнитной волны на плоскостойкой структуре из биизотропных материалов / В.Т. Ерофеенко, С.В. Малый // Информатика. – 2012. – № 1. – С. 58–65.
7. Ерофеенко, В.Т. Модели граничных условий на композиционных экранах для электромагнитных полей с осевой симметрией / В.Т. Ерофеенко // Известия НАН Беларуси. Сер. физ.-мат. наук. – 2010. – № 2. – С. 41–45.
8. Иванов, О.В. Распространение света в киральных слоистых средах. Метод матриц 4×4 / О.В. Иванов, Д.И. Семенов // Кристаллография. – 2000. – Т. 45, № 3. – С. 534–540.
9. Жук, Н.П. Скаляризация электромагнитного поля в произвольно-анизотропной плоскостойкой среде / Н.П. Жук // Радиотехника и электроника. – 1994. – № 11. – С. 1706–1715.
10. Евельсон, Р.Л. Метод решения задачи о распространении электромагнитных волн в многослойной плоскостойкой поглощающей анизотропной магнитоэлектрической среде / Р.Л. Евельсон // Радиотехника и электроника. – 2000. – Т. 45, № 6. – С. 918–925.

11. Шевченко, В.В. Локализация стационарного электромагнитного поля при помощи плоской границы метаматериала / В.В. Шевченко // Успехи физических наук. – 2011. – Т. 181, № 11. – С. 1171–1182.
12. Анютин, А.П. О резонансных свойствах ближнего поля в тонких линзах Веселаго и Пендри конечных размеров / А.П. Анютин, А.Д. Шатров // Радиотехника и электроника. – 2012. – Т. 57, № 9. – С. 1012–1019.
13. Ерофеенко, В.Т. Численное исследование взаимодействия электромагнитных полей электрического и магнитного диполей с композитным экраном / В.Т. Ерофеенко, В.Ф. Бондаренко // Известия НАН Беларуси. Сер. физ.-техн. наук. – 2013. – № 4. – С. 113–120.
14. Шевченко, В.В. Геометрооптическая теория плоской линзы из кирального метаматериала / В.В. Шевченко // Радиотехника и электроника. – 2009. – Т. 54, № 6. – С. 696–700.
15. Ерофеенко, В.Т. Двухсторонняя фокусировка электромагнитного поля точечного источника плоской линзой из метаматериала / В.Т. Ерофеенко // Информатика. – 2016. – № 1. – С. 14–25.
16. Шатров, А.Д. О разрешимости задач возбуждения плоскостойких сред из метаматериалов / А.Д. Шатров // Радиотехника и электроника. – 2007. – Т. 52, № 8. – С. 909–916.
17. Ерофеенко, В.Т. Фокусы по обе стороны плоской линзы из метаматериала при воздействии электромагнитного поля магнитного диполя / В.Т. Ерофеенко // Взаимодействие сверхвысокочастотного, терагерцевого и оптического излучения с полупроводниковыми микро- и наноструктурами, метаматериалами и биообъектами: материалы Всерос. науч. школы-семинара. – Саратов : Саратовский источник, 2016. – С. 69–71.

Поступила 06.08.2016

*Учреждение БГУ «НИИ прикладных проблем математики и информатики»,
Минск, пр. Независимости, 4
e-mail: bsu_erofeenko@tut.by*

V.T. Erofeenko

MULTIPLE FOCUSING OF ELECTROMAGNETIC FIELD OF MAGNETIC DIPOLE BY MEANS OF PLANE-PARALLEL TWO-LAYER LENS FROM METAMATERIALS

A method of analytical solution of the boundary-value problem of electromagnetic field penetration of point magnetic dipole through the infinite planar lens (screen) from the focusing metamaterials is suggested. It is demonstrated, that for the lenses with special combinations of material parameters in layers the series of focuses are formed: one is over the screen and three is in front of the screen in the half-space of point source of the excitation field. The summary field of focuses in front of the lens is equal to the sum of fields of magnetic and electric dipoles.

УДК 519.61/64:627.05

Н.В. Медвидь, П.Н. Мартынюк

РЕАЛИЗАЦИЯ АЛГОРИТМА БЕССЕТОЧНОГО МЕТОДА РАДИАЛЬНЫХ БАЗИСНЫХ ФУНКЦИЙ В ЗАДАЧАХ ПОДЗЕМНОЙ ГИДРОМЕХАНИКИ

Рассматривается математическая модель фильтрационной консолидации тела грунтовой плотины с водоводом и зоной размыва в двумерном случае, в которой учитывается влияние техногенных факторов (температуры и концентрации солей), а также проседание со временем верхней границы и внутренних точек плотины. Предлагается ПО для автоматизации расчетов численного решения краевой задачи бессеточным методом радиальных базисных функций, которое дает возможность проводить численные эксперименты, варьируя входные данные и параметры формы. На примере модельной задачи исследуется влияние наличия водовода и зоны размыва на распределение напоров и их градиентов, температуру и концентрацию солей в теле плотины на разных временных промежутках. Проводится ряд численных экспериментов и осуществляется их анализ.

Введение

Языки объектно-ориентированного программирования (ООП) как ядро объектно-ориентированной парадигмы программирования пользуются в последнее время большой популярностью среди разработчиков ПО. ООП – одна из парадигм, в соответствии с которой программа рассматривается как множество объектов, взаимодействующих между собой, и состоит из трех основных концепций: инкапсуляции, наследования и полиморфизма. Согласно парадигме ООП каждый объект способен получать сообщения, обрабатывать данные и отправлять сообщения другим объектам [1]. Идеи ООП, когда определенные математические структуры представляются классами и объектами, а операции над структурами (дифференцирование, интегрирование и т. д.) и взаимодействие между ними выражаются через методы классов, нашли широкое применение при программировании численных методов математической физики [2].

Не определяя своей целью классификацию численных методов математической физики, остановимся на двух больших группах этих методов: сеточных и бессеточных [3]. Численное решение краевых задач сеточными методами (например, методом конечных элементов) реализовано в специальных пакетах прикладных программ. Большинство из них могут работать на нескольких операционных системах. Поскольку пакетов прикладных программ по решению краевых задач бессеточными методами в настоящее время почти нет [4], публикация некоторых идей ООП для программной реализации алгоритмов одного из таких методов (на примере математической модели фильтрационной консолидации грунтовой плотины с водоводом) и является целью настоящей статьи.

Выбранный численный метод основывается на методе коллокаций в точках [5, 6] с использованием в качестве базисных радиальных функций [7, 8]. Первое историческое название метода – multiquadrics [9], позже – Kansa method [10]. Сейчас этот метод также называют radial basis function collocation method [11], radial basis function approximation method [12] или radial basis function method [13]. В настоящей работе применяют термин «метод радиальных базисных функций», т. е. метод коллокации в точках [5] с использованием в качестве базисных радиальных функций, хотя на данном этапе под этим термином подразумевается целый класс бессеточных методов математической физики, которые отличаются принципами построения расчетных схем (см., например, [14, 15]).

1. Математическая модель

При строительстве и эксплуатации плотин с водоводами остро стоит проблема их безаварийной работы [16]. Связана она с повреждениями в водоводах, размывом грунта с образованием пустотелых зон вдоль водовода, что может привести к полному разрушению плотин. Проблема осложняется еще больше, если плотина попадает в зону техногенных воздействий

(от химических растворов, изменений температуры, химической суффозии). Проведение натурных экспериментов – очень ресурсозатратный процесс. Поэтому для прогнозирования протекания подобных процессов целесообразно использовать математическое и компьютерное моделирование.

Математические модели взаимосвязанных процессов фильтрационной консолидации, фильтрационного разрушения грунтов и тепломассопереноса построены и рассмотрены, например, в работах [17–19]. Однако в этих работах рассмотрены только процессы контактного размыва грунтов и не затронуты процессы в грунтовых плотинах с инженерными включениями (водоводами), тем более с наличием зоны выноса (размыва) грунта через повреждение в водоводе. Применение метода радиальных базисных функций (РБФ) для численного решения математических моделей взаимосвязанных процессов фильтрационной консолидации и тепломассопереноса описано в работе [20], однако вопросы структур данных, используемых при практической реализации этого метода, там не затрагиваются.

Рассмотрим двухмерную задачу фильтрационной консолидации в теле грунтовой плотины, профиль которой изображен на рис. 1. Водовод разбивает плотину на две области: Ω_1 с границами $\Gamma_6, \Gamma_4, \Gamma_2, \Gamma_1$ и свободной границей Γ_0 и Ω_2 с границами $\Gamma_3, \Gamma_7, \Gamma_5, \Gamma_9$. Зона размыва – область Ω_3 с границами Γ_{12}, Γ_{11} . Здесь $\Gamma_4, \Gamma_7, \Gamma_2, \Gamma_9$ – границы контакта с водохранилищем; Γ_1 – граница контакта с атмосферой; $\Gamma_3, \Gamma_5, \Gamma_6, \Gamma_8, \Gamma_{10}$ – непроницаемые границы; Γ_{12} – граница, на которой выполняется условие быстрого выноса грунта; Γ_{11} – граница, на которой задаются условия сопряжения.

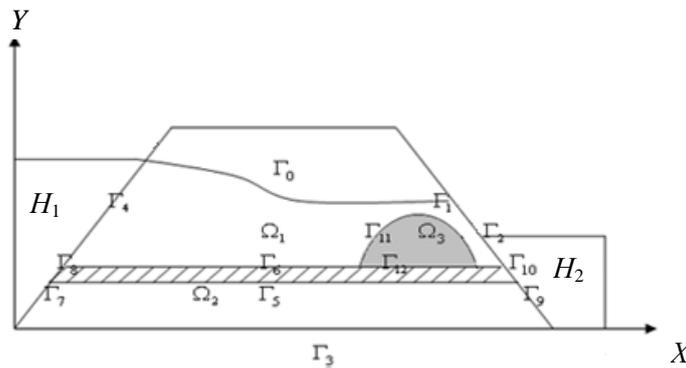


Рис. 1. Профиль тела грунтовой плотины с водоводом и зоной размыва

Сделаем два упрощающих предположения: зона размыва Ω_3 пока фиксированная и не меняет своего положения и формы во времени; для напоров в зоне размыва выполняется уравнение Лапласа.

Тогда математическая модель двухмерной задачи фильтрационной консолидации тела грунтовой плотины с водоводом и зоной размыва при учете ползучести скелета грунта и тепломассопереноса может быть описана краевой задачей, которая имеет следующий вид:

– уравнение фильтрационной консолидации

$$\begin{aligned}
 & 2a_0 \frac{\partial^2 h}{\partial t^2} + \beta_1(a_0 + a_1) \frac{\partial h}{\partial t} = \frac{(1+e)(1+\xi)}{\gamma} \beta_1 [\operatorname{div}(\mathbf{K}(\mathbf{C}, \mathbf{N}, \mathbf{T}) \nabla h) - \\
 & - \sum_{i=1}^n \operatorname{div}(\mathbf{K}_c \nabla c_i) - \operatorname{div}(\mathbf{K}_T \nabla \mathbf{T})] + \frac{(1+e)(1+\xi)}{\gamma} \frac{\partial}{\partial t} [\operatorname{div}(\mathbf{K}(\mathbf{C}, \mathbf{N}, \mathbf{T}) \nabla h) - \\
 & - \sum_{i=1}^n \operatorname{div}(\mathbf{K}_c \nabla c_i) - \operatorname{div}(\mathbf{K}_T \nabla \mathbf{T})] + \sum_{i=1}^n \frac{\varepsilon(1+e)(1+\xi)}{\rho_i \gamma} \times \\
 & \times \left[\beta_1 \left[\sigma \frac{\partial c_i}{\partial t} - e \frac{\partial N_i}{\partial t} \right] + \left[\sigma \frac{\partial^2 c_i}{\partial t^2} - e \frac{\partial^2 N_i}{\partial t^2} \right] \right], \quad \mathbf{X} \in \Omega_1 \cup \Omega_2, \quad t > 0;
 \end{aligned} \tag{1}$$

– уравнение фильтрации в зоне размыва (уравнение Лапласа)

$$\Delta h(\mathbf{X}, t) = 0, \mathbf{X} = (x, y) \in \Omega_3, t > 0; \quad (2)$$

– уравнения массопереноса

$$\operatorname{div}(D_T^{(i)} \nabla \mathbf{T}) + \operatorname{div}(D^{(i)} \nabla c_i) - (\mathbf{u}, \nabla c_i) = \sigma \frac{\partial c_i}{\partial t} + \frac{\partial N_i}{\partial t}, \quad (3)$$

$$\mathbf{X} \in \Omega_1 \cup \Omega_2, t > 0, i = \overline{1, n},$$

$$\operatorname{div}(D_T^{(i)} \nabla \mathbf{T}) + \operatorname{div}(D^{(i)} \nabla c_i) = 0, \mathbf{X} \in \Omega_3, t > 0, i = \overline{1, n}; \quad (4)$$

– уравнение кинетики массообмена

$$\frac{\partial N_i}{\partial t} = -\gamma_i (C_m^{(i)}(\mathbf{C}, \mathbf{N}, \mathbf{T}) - c_i) N_i^{\alpha_i}, \mathbf{X} \in \Omega_1 \cup \Omega_2, t > 0, i = \overline{1, n}; \quad (5)$$

– уравнения теплопереноса

$$\operatorname{div}(\lambda \nabla \mathbf{T}) - \rho c_p (\mathbf{u}, \nabla \mathbf{T}) = c_T \frac{\partial \mathbf{T}}{\partial t}, \mathbf{X} \in \Omega_1 \cup \Omega_2, t > 0, \quad (6)$$

$$\operatorname{div}(\lambda \nabla \mathbf{T}) = 0, \mathbf{X} \in \Omega_3, t > 0; \quad (7)$$

– обобщенный закон Дарси – Герсеванова

$$\mathbf{u} - e\mathbf{v} = -\mathbf{K}(\mathbf{C}, \mathbf{N}, \mathbf{T}) \nabla h + \sum_{i=1}^n \mathbf{K}_{c_i} \nabla c_i + \mathbf{K}_T \nabla \mathbf{T}, \mathbf{X} \in \overline{\Omega}, t \geq 0; \quad (8)$$

– начальные условия

$$h(\mathbf{X}, 0) = H_0(\mathbf{X}), \frac{\partial h(\mathbf{X}, 0)}{\partial t} = 0, \mathbf{X} \in \overline{\Omega}, \quad (9)$$

$$c_i(\mathbf{X}, 0) = C_0^{(i)}(\mathbf{X}), \mathbf{X} \in \overline{\Omega}, i = \overline{1, n}, \quad (10)$$

$$N_i(\mathbf{X}, 0) = N_i^{(0)}(\mathbf{X}), \mathbf{X} \in \overline{\Omega}, i = \overline{1, n}, \quad (11)$$

$$T(\mathbf{X}, 0) = T_0(\mathbf{X}), \mathbf{X} \in \overline{\Omega}; \quad (12)$$

– граничные условия и условия сопряжения

$$h(\mathbf{X}, t)|_{\Gamma_1 \cup \Gamma_0} = y, \mathbf{X} \in \Gamma_1 \cup \Gamma_0, h(\mathbf{X}, t)|_{\Gamma_2} = H_2(\mathbf{X}, t), \mathbf{X} \in \Gamma_2,$$

$$h(\mathbf{X}, t)|_{\Gamma_9} = H_2(\mathbf{X}, t), \mathbf{X} \in \Gamma_9, (\nabla h, \mathbf{n})|_{\Gamma_{10}} = 0, \mathbf{X} \in \Gamma_{10},$$

$$(\nabla h, \mathbf{n})|_{\Gamma_3} = 0, \mathbf{X} \in \Gamma_3, h(\mathbf{X}, t)|_{\Gamma_4} = H_1(\mathbf{X}, t), \mathbf{X} \in \Gamma_4,$$

$$h(\mathbf{X}, t)|_{\Gamma_7} = H_1(\mathbf{X}, t), \mathbf{X} \in \Gamma_7, (\nabla h, \mathbf{n})|_{\Gamma_3} = 0, \mathbf{X} \in \Gamma_3, \quad (13)$$

$$(\nabla h, \mathbf{n})|_{\Gamma_5} = 0, \mathbf{X} \in \Gamma_5, (\nabla h, \mathbf{n})|_{\Gamma_6} = 0, \mathbf{X} \in \Gamma_6,$$

$$h(x, y) = z, \mathbf{X} \in \Gamma_{12}, [h(x, y)] = 0, \mathbf{X} \in \Gamma_{11};$$

$$(\mathbf{D}^{(i)} \nabla c_i + \mathbf{D}_T^{(i)} \nabla T, \mathbf{n})|_{\Gamma_0 \cup \Gamma_1} = 0, \mathbf{X} \in \Gamma_0 \cup \Gamma_1, i = \overline{1, n},$$

$$c_i(\mathbf{X}, t)|_{\Gamma_2} = C_1^{(i)}(\mathbf{X}, t), \mathbf{X} \in \Gamma_2, i = \overline{1, n},$$

$$c_i(\mathbf{X}, t)|_{\Gamma_9} = C_1^{(i)}(\mathbf{X}, t), \mathbf{X} \in \Gamma_9, i = \overline{1, n}, (\nabla c_i, \mathbf{n})|_{\Gamma_{10}} = 0, \mathbf{X} \in \Gamma_{10}, i = \overline{1, n},$$

$$\begin{aligned} (\nabla c_i, \mathbf{n})|_{\Gamma_3} = 0, \mathbf{X} \in \Gamma_3, i = \overline{1, n}, c_i(\mathbf{X}, t)|_{\Gamma_4} = C_m^{(i)}(\mathbf{X}, t), \mathbf{X} \in \Gamma_4, i = \overline{1, n}, \\ c_i(\mathbf{X}, t)|_{\Gamma_7} = C_m^{(i)}(\mathbf{X}, t), \mathbf{X} \in \Gamma_7, i = \overline{1, n}, (\nabla c_i, \mathbf{n})|_{\Gamma_8} = 0, \mathbf{X} \in \Gamma_8, i = \overline{1, n}, \end{aligned} \quad (14)$$

$$(\nabla c_i, \mathbf{n})|_{\Gamma_5} = 0, \mathbf{X} \in \Gamma_5, i = \overline{1, n}, (\nabla c_i, \mathbf{n})|_{\Gamma_6} = 0, \mathbf{X} \in \Gamma_6, i = \overline{1, n},$$

$$\frac{\partial c_i}{\partial n} = 0, \mathbf{X} \in \Gamma_{12}, [c_i(x, y)] = 0, \mathbf{X} \in \Gamma_{11}, i = \overline{1, n};$$

$$T(\mathbf{X}, t)|_{\Gamma_0 \cup \Gamma_1 \cup \Gamma_2 \cup \Gamma_9} = T_{\text{atm}}(\mathbf{X}, t), \mathbf{X} \in \Gamma_0 \cup \Gamma_1 \cup \Gamma_2 \cup \Gamma_9,$$

$$(\nabla T, \mathbf{n})|_{\Gamma_3} = 0, \mathbf{X} \in \Gamma_3, T(\mathbf{X}, t)|_{\Gamma_4} = T_1(\mathbf{X}, t), \mathbf{X} \in \Gamma_4,$$

$$T(\mathbf{X}, t)|_{\Gamma_7} = T_1(\mathbf{X}, t), \mathbf{X} \in \Gamma_7, (\nabla T, \mathbf{n})|_{\Gamma_8} = 0, \mathbf{X} \in \Gamma_8,$$

$$(\nabla T, \mathbf{n})|_{\Gamma_{10}} = 0, \mathbf{X} \in \Gamma_{10}, (\nabla T, \mathbf{n})|_{\Gamma_5} = 0, \mathbf{X} \in \Gamma_5, (\nabla T, \mathbf{n})|_{\Gamma_6} = 0, \mathbf{X} \in \Gamma_6,$$

$$(\nabla T, n)|_{\Gamma_{12}} = 0, \mathbf{X} \in \Gamma_{12}, [T(x, y)] = 0, \mathbf{X} \in \Gamma_{11},$$

где $H_0(\mathbf{X}), H_1(\mathbf{X}, t), H_2(\mathbf{X}, t), C_0^{(i)}(\mathbf{X}), C_1^{(i)}(\mathbf{X}, t), C_m^{(i)}(\mathbf{X}, t), N_i^{(0)}(\mathbf{X}), T_0(\mathbf{X}), T_{\text{atm}}(\mathbf{X}, t), T_1(\mathbf{X}, t)$ – известные функции. Здесь приняты следующие обозначения: h – напор в поровой жидкости; c_i, N_i – концентрация многокомпонентного химического раствора в жидкой и твердой фазах; T – температура; \mathbf{u} – вектор скорости фильтрации; $\mathbf{K}(C, N, T)$ – тензор коэффициентов фильтрации; \mathbf{K}_{c_i} – тензор коэффициентов химического осмоса; \mathbf{K}_T – тензор коэффициентов термического осмоса; $\mathbf{D}^{(i)}, \mathbf{D}_T$ – тензоры коэффициентов конвективной диффузии и термодиффузии; λ – тензор коэффициентов эффективной теплопроводности влажного грунта; ρ – плотность порового раствора; c_p – удельная теплоемкость порового раствора; c_T – объемная теплоемкость грунта; C_m – концентрация предельного насыщения химических веществ в жидкой фазе; σ – пористость грунта; n – количество компонентов в солевом растворе; \mathbf{n} – вектор направляющих косинусов внешней нормали к рассматриваемой границе; a_0, a_1, β_1 – параметры ползучести скелета грунта (здесь аналогично [17] использованы элементы теории наследственной линейной ползучести); e – коэффициент пористости; ξ – коэффициент бокового давления грунта.

Свободная граница Γ_0 (кривая депрессии) описывается соотношением [21]

$$\varphi(\mathbf{X}, t) = (y - h(\mathbf{X}, t))|_{\Gamma_0} = 0, \quad (16)$$

и на ней выполняется кинематическое граничное условие [21]

$$(\mathbf{u}, \mathbf{n})|_{\Gamma_0} = \sigma \frac{\partial \varphi}{\partial t}. \quad (17)$$

Научная новизна математической модели (1)–(17) состоит в учете возможности фильтрационного разрушения тела грунтовой плотины через брешь в водоводе, а также в учете влияния техногенных факторов (температуры и многокомпонентных химических растворов). С физической точки зрения решение такой задачи требует рассмотрения взаимосвязанных процессов фильтрации и тепломассопереноса в теле грунтовой плотины (пористой среде) и в зоне размыва (области, которая заполнена суспензией). В научной литературе такие задачи еще не рассматривались. С математической точки зрения математическая модель физической задачи описыва-

ется начально-краевой задачей для системы эллиптических и параболических уравнений с условиями сопряжения на границах контакта. Такие задачи известны в научной литературе, однако их численное решение вызывает значительные трудности [21]. С деталями выведения уравнений модели (1)–(17) можно ознакомиться в работах [17–19] (см. также приведенную там библиографию).

Сформулированная задача является нелинейной краевой задачей для системы эллиптических и параболических уравнений в областях со свободными границами. Исследование подобных задач на корректность описано, к примеру, в работе [22] (см. также приведенную там библиографию). Однако в данной работе рассмотрен лишь одномерный случай и, кроме того, не учтено наличие зоны размыва. Поэтому в общем вопрос исследования корректности постановки указанной задачи с точки зрения классической математики остается открытым и, возможно, станет одним из направлений будущих исследований авторов.

2. Идея численного метода

Идея бессеточного метода РБФ (см., например, [3] и приведенную там библиографию по методу РБФ) заключается в следующем. Покроем замыкание области $\overline{\Omega} = \Omega \cup \Gamma$ узловыми точками (x_j, y_j) , $j = \overline{1, p}$ (некоторые из алгоритмов рассмотрены и реализованы в работах [23–25]). При этом любые две точки не должны совпадать. Далее покрываем замыкание $\overline{\Omega}$ области Ω коллокационными точками (x_i, y_i) , $i = \overline{1, m}$. Обозначим $\omega = \{i : (x_i, y_i) \in \Omega\}$, $\gamma_k = \{i : (x_i, y_i) \in \Gamma_k\}$, $k = \overline{0, 4}$. Введем обозначение $r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$, $j = \overline{1, p}$, $i = \overline{1, m}$, $m \geq p$. Приближенные значения функций будем искать в виде

$$z(\mathbf{X}, t) \approx \sum_{j=1}^n z_j(t) \varphi_j(r_j, \varepsilon_z),$$

где ε_z – параметр формы; $\varphi_j = \varphi(r_j, \varepsilon_z)$ – радиальная базисная

функция, зависящая от расстояния $r_j = \sqrt{(x - x_j)^2 + (y - y_j)^2}$, $j = \overline{1, p}$, и параметра формы $\varepsilon_z > 0$; $z_j(t)$, $j = \overline{1, p}$, – неизвестные коэффициенты для нахождения значений соответственно функций напора, концентрации солей в жидкой и твердой фазах и температуры, которые зависят только от времени.

В результате использования метода коллокации в точках [5] для отыскания неизвестных коэффициентов получим задачу Коши для системы обыкновенных нелинейных дифференциальных уравнений. Ввиду громоздкости выкладок и в связи с другими целями этой статьи не будем приводить здесь вид этой задачи. Более детальную информацию читатель может получить, например, в работе [3]. Отметим только, что в научной литературе исследование точности метода РБФ проведено лишь для некоторых краевых задач (см., например, [7, 8, 26, 27]). В настоящее время единственным инструментом исследования рассмотренных в статье краевых задач является вычислительный эксперимент – сравнение решений метода РБФ с решениями, найденными другими методами, которые хорошо теоретически обоснованы (например, методом конечных элементов), как это сделано в работе [3].

3. Структура и описание основных классов

Программное приложение разработано в интегрированной среде разработки Microsoft Visual Studio 2008 на языке C#. Преимуществом программного приложения является использование стандартных средств и функций для создания графиков. Поэтому от пользователя не требуется установки DirectX- или OpenGL-библиотек. Для удобства разработки и дальнейшей поддержки программного приложения все составляющие классы были разделены на четыре основные категории (рис. 2 и 3).

В группу Interfaces входят классы, которые служат прототипами для входных параметров некоторых функций и обеспечивают совместимость между узловыми и коллокационными точками при работе с определенными процедурами. На данный момент там находится только один класс IPoint, что является обобщенным интерфейсом точки в пространстве.

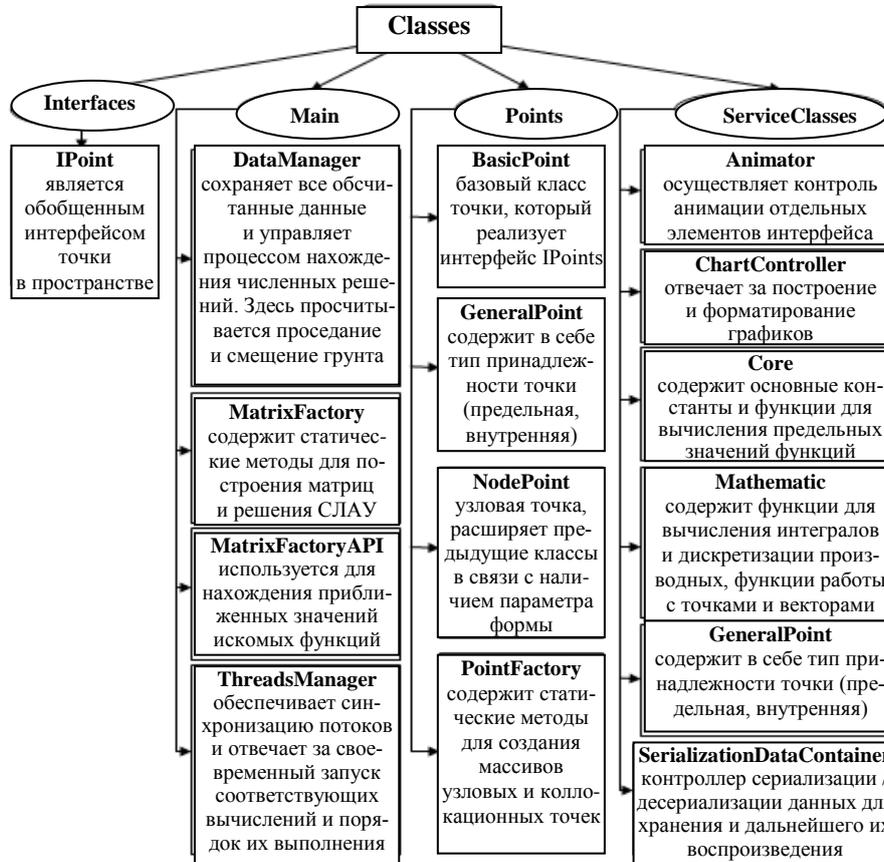


Рис. 2. Структура основных классов

Следующей по важности является группа ServiceClasses. Классы этой группы используются для упрощения вычислений и выполнения технических операций (сохранения, восстановления данных, построения графиков), а также сохранения входных данных программного приложения.

Затем следует группа классов под названием Points. Как можно догадаться из названия, основным назначением классов этой группы является описание и организация работы с узловыми и коллокационными точками.

Главной является группа Main. Она включает классы, используемые непосредственно для отыскания решений.

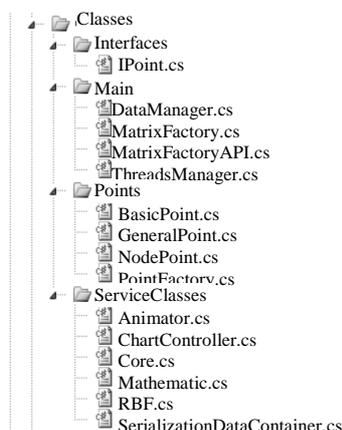


Рис. 3. Иерархия классов программного приложения

После запуска программного приложения перед пользователем появляется главное окно программы (рис. 4).

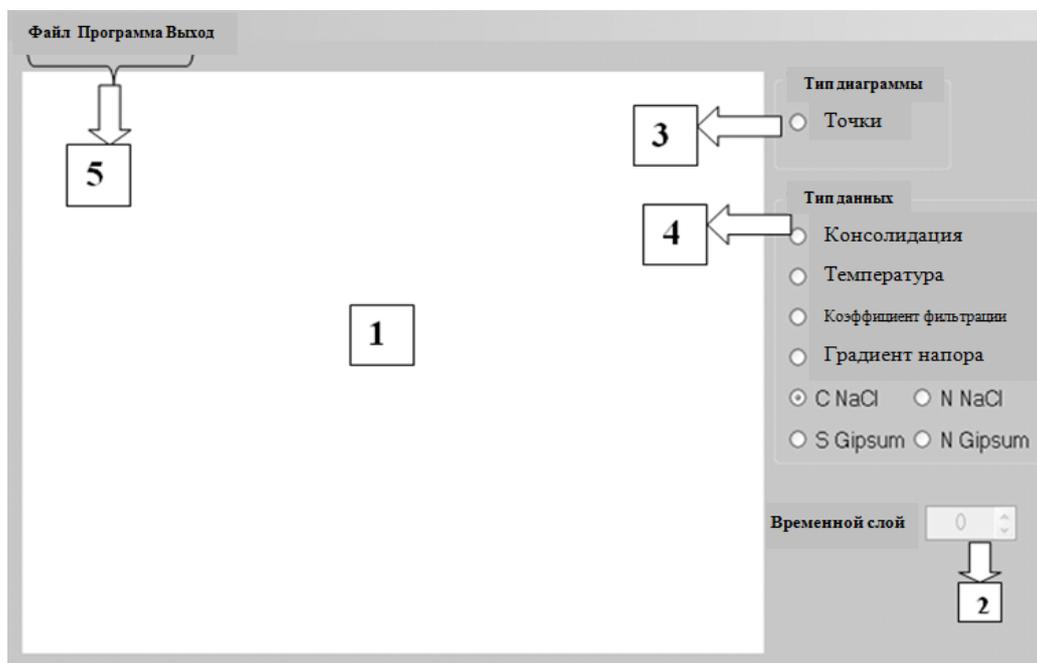


Рис. 4. Главное окно программы

Интерфейс пользователя можно разделить на семь основных частей:

- главное полотно для построения графиков в ходе выполнения программы;
- переключатель временных промежутков (если процесс расчета не запущен, то данный элемент выключен);
- секцию, в которой выбирается способ отображения графиков (доступной является возможность отображения только точками);
- секцию для выбора графика функции (избыточных напоров, концентрации соли и гипса в твердой и жидкой фазах, температуры, градиента напора, а также графика аппроксимации коэффициента фильтрации);
- главное меню для запуска вычислений, сохранения результатов работы, выхода, а также загрузки сохраненных результатов;
- панель управления параметрами формы (в свернутом виде);
- панель для отражения внешних нормалей к поверхности, что позволяет динамично наблюдать за их изменением в процессе работы программы (в свернутом виде).

После запуска программы у пользователя есть два возможных алгоритма работы с ней:

- нажать меню Программа-Запуск, программа начнет вычислять данные и покажет соответствующие графики. После того как все расчеты будут завершены, пользователю будет доступна возможность сохранения данных вычислений в файл (меню Файл-Сохранить станет активным). При этом имя файла генерируется автоматически, исходя из входных данных и количества точек;
- загрузить готовые данные (нажать на меню Файл-Открыть), после чего появится окно выбора файлов для загрузки в программное приложение (рис. 5).

Окно загрузки сохраненных данных содержит два текстовых поля, где после выбора файлов отображается их название. Чтобы загрузить файл, нужно нажать на выделенную кнопку справа от этого поля, после чего можно будет указать файл на диске. Обычно окно позволяет загрузить только один файл. Однако если включить соответствующий флажок (Загрузить два файла), то кнопка для загрузки второго файла станет активной и в результате программа отобразит разницу (в математическом смысле) между выбранными данными. Цель подсчета разни-

цы – отображение, например, разницы распределения избыточных напоров в случаях с учетом влияния теплосолепереноса и без учета.

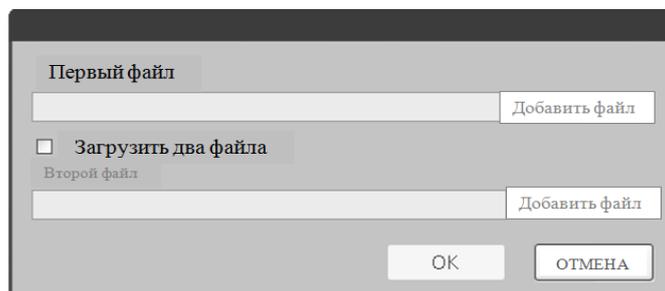


Рис. 5. Окно загрузки сохраненных данных

4. Схема отыскания координат точек свободной границы

Свободная граница области будет со временем смещаться, т. е. координаты точек на границе Γ_0 будут изменяться. Поэтому необходим алгоритм отыскания этих координат в соответствующий момент времени.

Дискретизируем уравнение (17) по времени, получим

$$(\mathbf{u}^{(k)}, \mathbf{n}^{(k)}) = \sigma \frac{\varphi^{(k+1)}(\mathbf{X}) - \varphi^{(k)}(\mathbf{X})}{\tau},$$

где $(\varphi^{(k+1)} - \varphi^{(k)})$ – смещение свободной границы на k -м временном слое в направлении внешней нормали. Далее имеем

$$\varphi^{(k+1)}(\mathbf{X}) - \varphi^{(k)}(\mathbf{X}) = \frac{\tau}{\sigma} (u_x^{(k)}(\mathbf{X})n_x^{(k)}(\mathbf{X}) + u_y^{(k)}(\mathbf{X})n_y^{(k)}(\mathbf{X})), \quad \mathbf{X} \in \Gamma_0. \quad (18)$$

Здесь $f^{(k)}(\mathbf{X}) = f(\mathbf{X}, t_k)$, $t_k = \tau \cdot k$, $k = 0, 1, 2, \dots$, где τ – шаг дискретизации по времени; $\mathbf{u} = (u_x(X, t), u_y(X, t))$ – вектор скорости фильтрации; $\mathbf{n} = (n_x(X, t), n_y(X, t))$ – вектор направляющих косинусов внешней нормали к рассматриваемой границе.

5. Алгоритм нахождения координат внутренних точек области

Поскольку при смещении верхней границы некоторые внутренние точки области могут оказаться вне области, необходимо реализовать алгоритм уплотнения внутренних точек. Для этой цели подходит алгоритм уплотнения точек на величину, которая является обратно пропорциональной квадрату расстояния от внутренней точки до ближайшей точки свободной границы Γ_0 :

$$\mathbf{X}_i^{(k+1)} = \mathbf{X}_i^{(k)} + \frac{(\varphi_j^{(k+1)} - \varphi_j^{(k)})}{r(\mathbf{X}_i^{(k)}, \mathbf{X}_j^{(k)})^2}, \quad i \in \omega, \quad (19)$$

где $r(\mathbf{X}_i^{(k)}, \mathbf{X}_j^{(k)}) = \min_{p \in \gamma_0} (r(\mathbf{X}_i^{(k)}, \mathbf{X}_p^{(k)}))$, $j \in \gamma_0$, $r(\mathbf{X}_i^{(k)}, \mathbf{X}_j^{(k)}) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$.

6. Результаты численного эксперимента

Рассмотрим двухмерную модельную задачу фильтрационной консолидации грунтовой плотины высотой 20 м и длиной нижнего основания 80 м при наличии водовода (рис. 6).

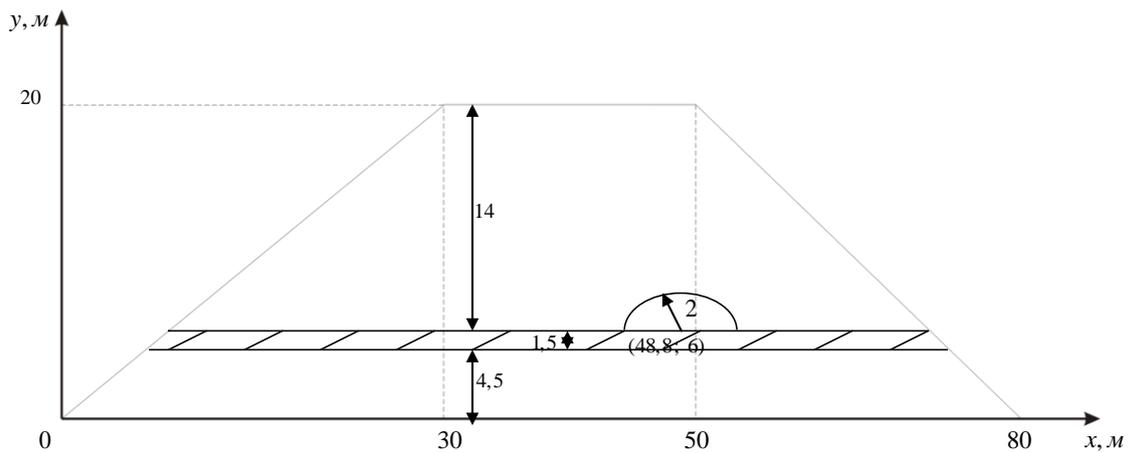


Рис. 6. Схема и размеры плотины в модельных расчетах

Входные данные задачи:

$$n = 0,4, \quad e = 0,7, \quad \alpha = 0,5, \quad \beta = 0,5, \quad a_0 = 4 \cdot 10^{-7} \text{ м}^2 / \text{Н}, \quad a_1 = 2 \cdot 10^{-6} \text{ м}^2 / \text{Н},$$

$$\rho_C = 2100 \text{ кг} / \text{м}^3, \quad \rho_S = 2300 \text{ кг} / \text{м}^3, \quad C_m = 160 \text{ г} / \text{л},$$

$$D_{C11} = D_{C22} = 0,08 \text{ м}^2 / \text{сут}, \quad D_{C12} = D_{C21} = 0 \text{ м}^2 / \text{сут},$$

$$D_{S11} = D_{S22} = 0,008 \text{ м}^2 / \text{сут}, \quad D_{S12} = D_{S21} = 0 \text{ м}^2 / \text{сут},$$

$$K_C = 6 \cdot 10^{-7} \text{ м}^5 / \text{кг} \cdot \text{сут},$$

$$K_{C11} = K_{C22} = 6 \cdot 10^{-7} \text{ м}^2 / \text{кг} \cdot \text{сут}, \quad K_{C12} = K_{C21} = 0 \text{ м}^2 / \text{кг} \cdot \text{сут},$$

$$K_{S11} = K_{S22} = 6 \cdot 10^{-8} \text{ м}^2 / \text{кг} \cdot \text{сут}, \quad K_{S12} = K_{S21} = 0 \text{ м}^2 / \text{кг} \cdot \text{сут},$$

$$K_{T11} = K_{T22} = 6 \cdot 10^{-8} \text{ м}^2 / \text{кг} \cdot \text{сут}, \quad K_{T12} = K_{T21} = 0 \text{ м}^2 / \text{кг} \cdot \text{сут},$$

$$\gamma_S = 6,5 \cdot 10^{-5} \text{ сут}^{-1}, \quad \gamma_C = 6,5 \cdot 10^{-5} \text{ сут}^{-1}, \quad \gamma = 1,1 \cdot 10^4 \text{ кг} / (\text{м}^2 \cdot \text{сут}^2),$$

$$c_T = 2137 \text{ кДж} / \text{м}^3 \cdot \text{град}, \quad c_p = 4,2 \text{ кДж} / \text{кг} \cdot \text{град},$$

$$\lambda_{11} = \lambda_{22} = 108 \text{ кДж} / \text{м} \cdot \text{град} \cdot \text{сут}, \quad \lambda_{12} = \lambda_{21} = 0 \text{ кДж} / \text{м} \cdot \text{град} \cdot \text{сут},$$

$$T_0(x) = 4 \text{ }^\circ\text{C}, \quad T_{\text{атм}} = 17 + 13 \cos\left(\frac{\pi t}{180}\right),$$

$$C_0(x) = 8 \text{ г} / \text{л}, \quad S_0(x) = 0,1 \text{ г} / \text{л},$$

$$C_1(t) = 8 \text{ г} / \text{л}, \quad S_1(t) = 0,1 \text{ г} / \text{л},$$

$$H_0 = \sqrt{(H_1)^2 - \frac{(H_1)^2 - (H_2)^2}{L} x}, \quad H_1(t) = 20 \text{ м}, \quad H_2(t) = 10 \text{ м}.$$

Для численных экспериментов использовалась мультиквадратическая радиальная базисная функция $\varphi(r, \varepsilon) = \sqrt{1 + (r\varepsilon)^2}$; количество узловых (и коллокационных) точек равно 293. В области грунта использовалось значение параметра формы $\varepsilon = 1,3$, а в области размыва – $\varepsilon = 100$. Выбор значений параметров формы осуществлялся путем вычислительного эксперимента из условий устойчивости численных решений. Известные рекомендации по выбору параметров формы в этом численном примере не использовались. Процесс исследовался в течение 360 дней с шагом по времени 10 суток. Грунт считался изотропным и однородным.

Получены значения концентрации соли, гипса и напоров на момент времени 360 дней, которые показаны графически на рис. 7–9.

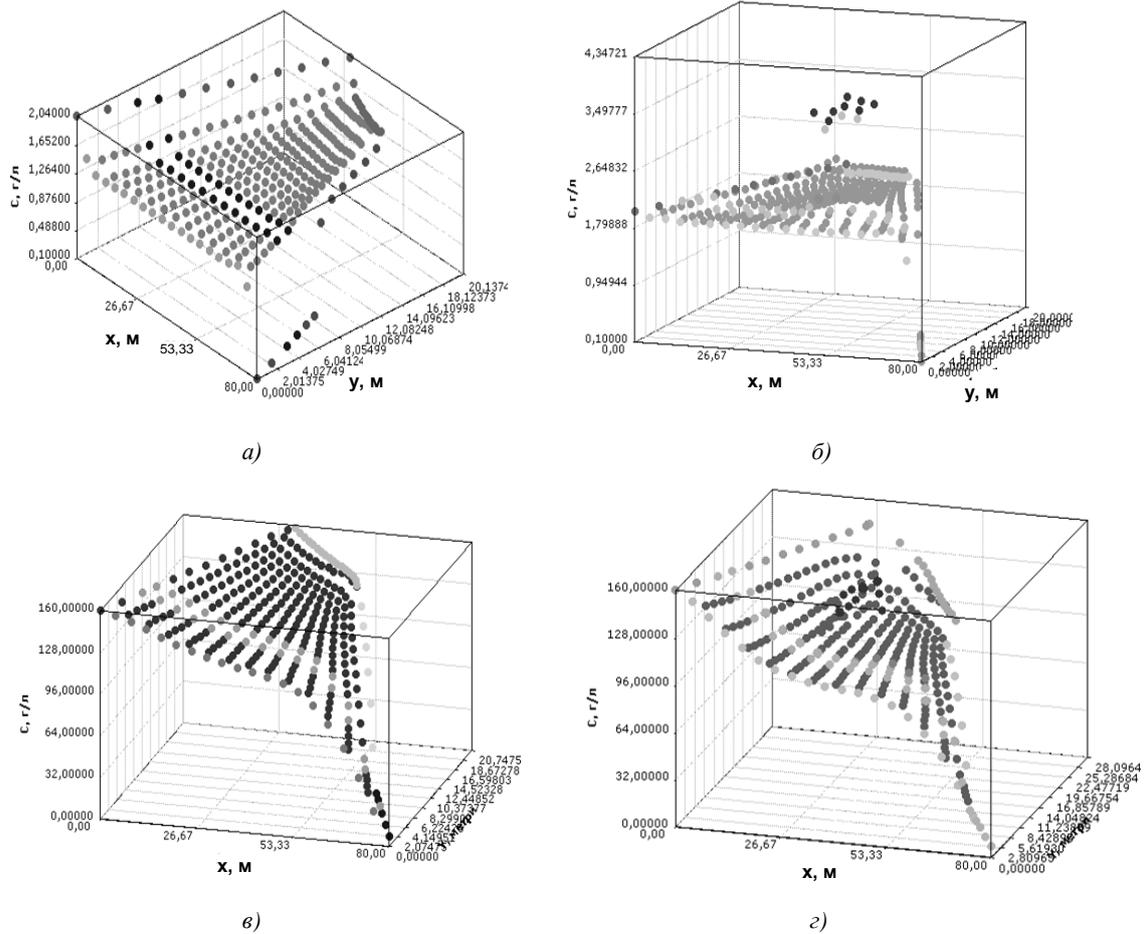


Рис. 7. Распределение концентрации: а) гипса в грунтовой плотине; б) гипса в грунтовой плотине с водоводом и зоной размыва; в) соли в грунтовой плотине; г) соли в грунтовой плотине с водоводом и зоной размыва

Наличие зоны размыва в теле плотины не оказывает значительного влияния на распределение концентрации гипса (рис. 7, а, б) и на концентрацию соли в жидкой фазе (рис. 7, в, г)

Численные эксперименты подтверждают, что наличие зоны размыва в теле плотины оказывает существенное влияние на распределение избыточных напоров, особенно в самой зоне размыва, где напор становится значительно меньше (рис. 8).

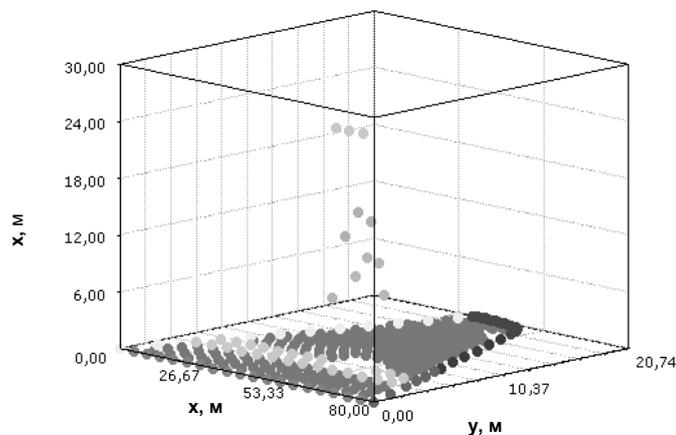


Рис. 8. Разница избыточных напоров в грунтовой плотине с водоводом и с водоводом и зоной размыва

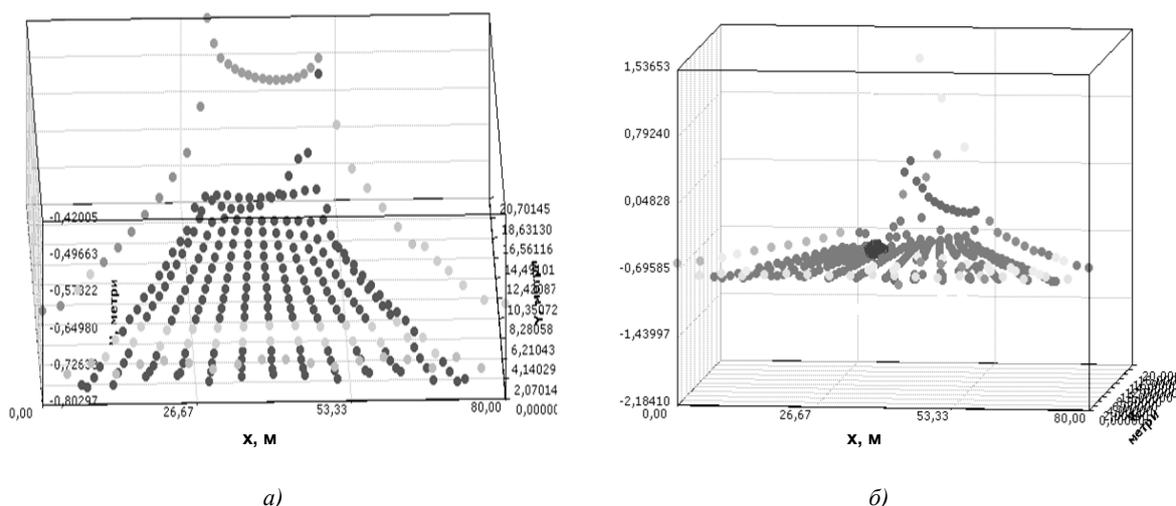


Рис. 9. Распределение компоненты $\partial h / \partial y$ градиента напоров: а) в грунтовой плотине; б) в грунтовой плотине с водоводом и зоной размыва

Эксперименты показывают, что наличие зоны размыва в теле плотины при приведенных входных данных значительным образом влияет на градиент напора (рис. 9). Известно, что увеличение градиента напора является одним из условий интенсификации процессов размыва грунтов фильтрационным потоком [18]. В зоне размыва он становится значительно больше (максимальная разница около двух). Максимальная относительная разница между распределением избыточных напоров в теле плотины только с водоводом и водоводом и зоной размыва соответственно для этой модельной задачи составляет около 71 %.

Заключение

Математическое моделирование и исследование процессов фильтрационной консолидации в теле грунтовой плотины приводит к сложным нелинейным краевым задачам математической физики. Поэтому необходимо разрабатывать численные методы, которые позволяют решать задачи такого типа.

Объектно-ориентированные языки программирования как средство реализации алгоритмов решения краевых задач являются весьма востребованными. Они дают возможность многократного использования текстов программ. Еще одно их преимущество заключается в способе комплектования объектно-ориентированных программных модулей. Благодаря этим преимуществам, а также, например, организации параллельных вычислений ООП является в настоящее время наиболее перспективным, распространенным и эффективным направлением в программировании.

Бессеточные методы решения краевых задач позволяют применять для дискретизации только узловую сетку, что значительно уменьшает количество использованных ресурсов и времени. Разработанное программное приложение дает возможность проводить многочисленные эксперименты, варьируя исходные данные и параметры формы, а также отображать результаты этих экспериментов в виде графиков и численных данных соответственно на каждом временном слое.

Суммируя полученные результаты относительно рассмотренного модельного примера, следует отметить, что наличие зоны размыва приводит к резкому возрастанию градиентов напоров в данной зоне. Такой рост градиентов напоров ведет к увеличению скорости движения поровой жидкости и, как следствие, к увеличению зоны размыва и возможному разрушению тела плотины. Пренебрегать возможными зонами размыва нельзя, и в случае проектирования реальных гидротехнических объектов необходимо проводить предварительное компьютерное прогнозирование таких процессов.

Список литературы

1. Kindler, E. Object-oriented simulation of systems with sophisticated control / E. Kindler, I. Krivy // *Intern. J. of General Systems*. – 2011. – Vol. 40. – P. 313–343.
2. Корчевская, Е.А. Современные численные методы в объектно-ориентированном изложении / Е.А. Корчевская // *Наука – образованию, производству, экономике : материалы XX(67) Региональной науч.-практ. конф. преподавателей, науч. сотрудников и аспирантов, Витебск*. – Витебск : ВГУ им. П.М. Машерова, 2015. – Т. 1. – С. 10–11.
3. Мартынюк, П.Н. Решение краевых задач для систем квазилинейных параболических уравнений сеточными и бессеточными численными методами / П.Н. Мартынюк // *Сибирские электронные математические известия [Электронный ресурс]*. – 2014. – Т. 11. – С. 476–493. – Режим доступа : <http://semr.math.nsc.ru/v11/p476-493.pdf>. – Дата доступа : 18.05.2016.
4. Liu, G.R. Meshfree methods. Moving beyond the finite element method / G.R. Liu. – CRC Press, 2010. – 792 p.
5. Власова, Е.А. Приближенные методы математической физики / Е.А. Власова, В.С. Зарубин, Г.Н. Кувырнин. – М. : Изд-во МГТУ им. Н.Э. Баумана, 2001. – 700 с.
6. Liu, G.R. An introduction to meshfree methods and their programming / G.R. Liu, Y.T. Gu. – Springer, 2005. – 480 p.
7. Franke, C. Solving partial differential equations by collocation using radial basis function / C. Franke, R. Schaback // *Appl. Math. Comp.* – 1998. – Vol. 93. – P. 73–82.
8. Franke, C. Convergence order estimates of meshless collocation methods using radial basis functions / C. Franke, R. Schaback // *Advances in computational mathematics*. – 1998. – Vol. 8(4). – P. 381–399.
9. Kansa, E.J. Multiquadrics – a scattered data approximation scheme with applications to computational fluid-dynamics. II. Solutions to parabolic, hyperbolic and elliptic partial differential equations / E.J. Kansa // *Comput. Math. Appl.* – 1990. – Vol. 19. – P. 147–161.
10. Pang, G. Space-fractional advection-dispersion equations by the Kansa method / G. Pang, W. Chen, Z. Fu // *J. of Computational Physics*. – 2015. – Vol. 293. – P. 280–296.
11. Mavric, B. Local radial basis function collocation method for linear thermoelasticity in two dimensions / B. Mavric, B. Sarler // *Intern. J. Numerical Methods for Heat and Fluid Flow*. – 2015. – Vol. 25, no. 6. – P. 148–1510.
12. Uddin, M. On the selection of a good value of shape parameter in solving time-dependent partial differential equations using RBF approximation method / M. Uddin // *Applied Mathematical Modelling*. – 2014. – Vol. 38. – P. 135–144.
13. Pearson, J.W. A radial basis function method for solving PDE constrained optimization problems / J.W. Pearson // *Numerical Algorithms*. – 2013. – Vol. 64. – P. 481–506.
14. Толстых, А.И. Бессеточный метод на основе радиальных базисных функций / А.И. Толстых, Д.А. Ширококов // *Журнал вычисл. матем. и матем. физ.* – 2005. – Т. 45, № 8. – С. 1498–1505.
15. Biazar, J. Galerkin RBF for integro-differential equations / J. Biazar, M.A. Asadi // *British Journal of Mathematics and Computer Science*. – 2015. – Vol. 11(2). – P. 1–9.
16. Анахаев, К.Н. Об авариях и повреждениях земляных плотин с водоводами: причины и способы совершенствования противотрационной защиты / К.Н. Анахаев, К.А. Гегиев, Б.Х. Амшоев // *Гидротехническое строительство*. – 2014. – № 3. – С. 30–36.
17. Власюк, А.П. Фильтрационная консолидация трехфазных грунтов с учетом ползучести скелета и влияния солепереноса в неизотермическом режиме / А.П. Власюк, П.Н. Мартынюк // *Математическое моделирование*. – 2010. – Т. 22, № 4. – С. 32–56.
18. Власюк, А.П. Контактный размыв и фильтрационная консолидация грунтов в условиях теплосолепереноса / А.П. Власюк, П.Н. Мартынюк // *Математическое моделирование*. – 2012. – Т. 24, № 11. – С. 97–112.
19. Мичута, О.Р. Моделирование влияния химической суффозии на фильтрационную консолидацию засоленных грунтов в неизотермических условиях / О.Р. Мичута, А.П. Власюк, П.Н. Мартынюк // *Математическое моделирование*. – 2013. – Т. 25, № 2. – С. 3–18.

20. Vlasyuk, A.P. Numerical solution of three-dimensional problems of filtration consolidation with regard for the influence of technogenic factors by the method of radial basis functions / A.P. Vlasyuk, P.M. Martynyuk // J. of Mathematical Sciences. – 2010. – Vol. 171, no. 5. – P. 632–648.

21. Сергиенко, И.В. Математическое моделирование и исследование процессов в неоднородных средах / И.В. Сергиенко, В.В. Скопецкий, В.С. Дейнека. – Киев : Наук. думка, 1991. – 432 с.

22. Martynyuk, P.M. Existence and uniqueness of a solution of the problem with free boundary in the theory of filtration consolidation of soil with regard for the influence of technogenic factors / P.M. Martynyuk // J. of Mathematical Sciences. – 2015. – Vol. 207, no. 1. – P. 59–73.

23. Мартынюк, П.Н. Узловое и конечно-элементное покрытия двумерных областей: некоторые алгоритмы и их программная реализация / П.Н. Мартынюк, Ю.А. Семенчук // Вестник Нац. ун-та водн. хоз-ва и природопользования. Сер. «Технические науки». – 2010. – Вып. 4(52). – С. 202–209.

24. Du, Q. Meshfree, probabilistic determination of point sets and support regions for meshless computing / Q. Du, M. Gunzburger, L. Ju // Computer Methods in Applied Mechanics and Engineering. – 2002. – Vol. 191(13–14). – P. 1349–1366.

25. Fornberg, B. Fast generation of 2-D node distributions for mesh-free PDE discretizations / B. Fornberg, N. Flyer // Computers and Mathematics with Applications. – 2015. – Vol. 69(7). – P. 531–544.

26. Lee, C.-F. On convergent numerical algorithms for nonsymmetric collocation / C.-F. Lee, L. Ling, R. Schaback // Advances in computational mathematics. – 2009. – Vol. 30(4). – P. 339–354.

27. Ling, L. Stable and convergent nonsymmetric meshless collocation methods / L. Ling, R. Schaback // SYAM J. on numerical analysis. – 2008. – Vol. 46(3). – P. 1097–1115.

Поступила 03.08.2016

*Национальный университет водного
хозяйства и природопользования,
Украина, г. Ровно, ул. Соборная, 11
e-mail: natasha_medvid@mail.ru,
martinjuk@ukr.net*

N.V. Medvid, P.M. Martynyuk

THE ALGORITHM OF MESHFREE METHOD OF RADIAL BASIS FUNCTIONS IN TASKS OF UNDERGROUND HYDROMECHANICS

A Mathematical model of filtering consolidation in the body of soil dam with conduit and washout zone in two-dimensional case is considered. The impact of such technogenic factors as temperature, salt concentration, subsidence of upper boundary and interior points of the dam with time is taken into account. The software to automate the calculation of numerical solution of the boundary problem by radial basis functions has been created, which enables to conduct numerical experiments by varying the input parameters and shape. The influence of the presence of conduit and washout zone on the pressure, temperature and concentration of salts in the dam body at different time intervals is investigated. A number of numerical experiments is conducted and the analysis of dam accidents is performed.

УДК 004.942: 537.87

К.С. Курочка

КОНЕЧНО-ЭЛЕМЕНТНОЕ МОДЕЛИРОВАНИЕ РАСПРЕДЕЛЕНИЯ ЭЛЕКТРОМАГНИТНОГО ПОЛЯ В БЛИЖНЕЙ ЗОНЕ СФЕРИЧЕСКОЙ НАНОЧАСТИЦЫ

Предлагается математическая модель распределения электромагнитного поля, вызываемого плоской волной в ближней зоне металлической или диэлектрической сферической наночастицы, на основе метода конечных элементов, который позволяет учесть неоднородную структуру расчетной области. Разработаны соответствующие алгоритмы и программное обеспечение, проведена их верификация. Расхождение результатов исследования предлагаемой математической модели и имеющихся результатов расчетов по аналитическим формулам согласно теории Ми (G. Mie) [1] для однородных сферических наночастиц не превышает 10 %.

Введение

Материалы, представляющие собой пленки и стекла, которые содержат металлические наночастицы, находят или могут найти применение при создании таких объектов, как среды для регистрации оптических сигналов, в том числе быстропротекающих, оптические сенсорные элементы для селективного детектирования газообразных и (или) жидких веществ, преобразователи и (или) усилители излучения, в том числе вторичного, спектральные фильтры и покрытия, управляющие оптические и (или) оптоэлектронные элементы [2, 3].

Получение новых нанокompозитов, удовлетворяющих заданным свойствам и характеристикам, на основе наночастиц восстановленных металлов сопряжено с проведением большого числа дорогостоящих натуральных экспериментов [2, 4]. Применение инструментария компьютерного математического моделирования для исследования взаимодействия оптического излучения с композиционными наноматериалами различной внутренней структуры позволяет существенно снизить количество проводимых натуральных экспериментов и, как следствие, снизить затраты и время на получение новых материалов.

Проведение моделирования базируется на решениях системы уравнений Максвелла (J.C. Maxwell) [1, 4]. В настоящее время существует достаточное количество различных методов и алгоритмов решения указанной задачи [4–7]. Однако исследование математической модели сопряжено с многократным решением одной и той же задачи с различным набором входных параметров, что накладывает жесткие требования на алгоритмы как по времени нахождения единичного решения, так и по степени его адекватности исходной физической системе на всем диапазоне изменения значений входных параметров [7].

Дополнительные сложности вызывает то, что материалы на основе диоксида кремния с наночастицами восстановленных металлов являются неоднородными, причем эта неоднородность в большей степени проявляется в структуре самой металлической наночастицы [2], отказ от учета которой может привести к погрешностям при определении значений электромагнитного поля в ближней зоне.

В целях учета неоднородной структуры металлической наночастицы и композиционного материала был выбран метод конечных элементов (МКЭ) [8, 9].

1. Постановка задачи

Рассмотрим задачу определения интенсивности электромагнитного поля в ближней зоне сферической наночастицы. На частицу падает плоская монохроматическая волна. Начало координат поместим в центр частицы. Расчетная область содержит частицу, «ближнюю» и «дальнюю» зоны частицы и полностью поглощающий слой (рис. 1). В качестве граничных условий используются условия поглощения (ABC – absorbing boundary condition) [6, 10], реализованные в виде условия излучения Зоммерфельда (A. Sommerfeld) [6, 8].

«Ближняя» зона частицы (область 2, рис. 1) представляет собой параллелепипед, описывающий сферу радиуса $R + 2\lambda$, где R – радиус наночастицы, λ – длина падающей волны. Размеры «дальней» зоны определяются исходя из реализации граничных условий поглощения и представляют собой также параллелепипед с длиной ребра $7R$.

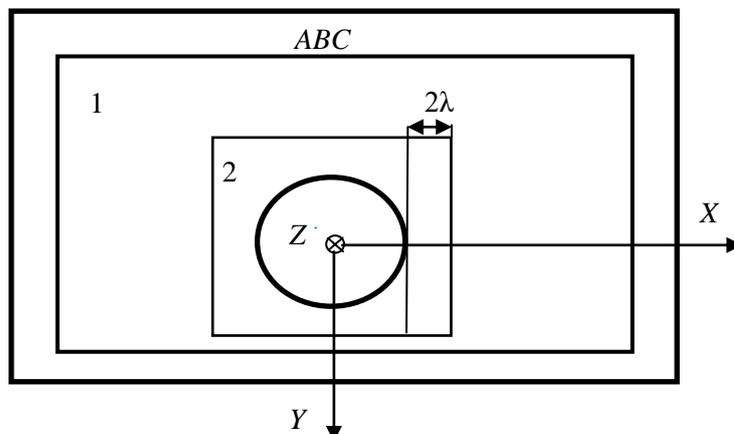


Рис. 1. Схема расчетной области (сечение расчетной 3D-области в координатной плоскости XY)

Напряженность электрического поля, создаваемого источником с плотностью зарядов \mathbf{J}_{imp} в области Ω , которая характеризуется электрической ϵ и магнитной μ постоянными, может быть описана с помощью уравнений Максвелла [1, 4]:

$$\nabla \times \mathbf{E} = -i\omega\mu\mathbf{H}; \quad (1)$$

$$\nabla \times \mathbf{H} = i\omega\epsilon\mathbf{E} + \mathbf{J}_{\text{imp}}; \quad (2)$$

$$\nabla \cdot (\epsilon\mathbf{E}) = -\frac{1}{i\omega} \nabla \cdot \mathbf{J}_{\text{imp}}; \quad (3)$$

$$\nabla \cdot (\mu\mathbf{H}) = 0, \quad (4)$$

где \mathbf{E} – напряженность электрического поля; \mathbf{H} – напряженность магнитного поля; ω – частота источника; i – мнимая единица; ∇ – гамильтонов (W.R. Hamilton) оператор.

Исследуемая область может быть как двух-, так и трехмерной. Для определения напряженности электрического поля \mathbf{E} необходимо решить уравнения Максвелла (1)–(4) с учетом граничных условий.

Типовые граничные условия для электрических полей включают однородные условия Дирихле (P.G.L. Dirichlet) на идеально проводящей поверхности, а также смешанные условия на границе поверхности, обладающей волновым сопротивлением. Формулировку этих граничных условий можно записать в виде

$$\mathbf{n} \times \mathbf{E} = \mathbf{P}; \quad (5)$$

$$\mathbf{n} \times \left(\frac{1}{\mu_r} \nabla \times \mathbf{E} \right) + \frac{ik_0}{\eta_r} \mathbf{n} \times (\mathbf{n} \times \mathbf{E}) = \mathbf{K}_N, \quad (6)$$

где \mathbf{n} – единичный вектор нормали к границам областей и подобластей; \mathbf{P} – установленные значения для компонент поля на Γ_D ; η_r – нормальный импеданс поверхности на Γ_N ; \mathbf{K}_N – известные функции, описанные на границе источника.

За счет исключения напряженности магнитного поля \mathbf{H} в (1) и преобразования (2) можно получить волновое уравнение, называемое уравнением Гельмгольца (H. Helmholtz) [6, 8]:

$$\nabla \times \left(\frac{1}{\mu_r} \nabla \times \mathbf{E} \right) - k_0^2 \varepsilon_r \mathbf{E} = -ik_0 Z_0 \mathbf{J}_{\text{imp}} \text{ на } \Omega, \quad (7)$$

где $\mu_r = \mu/\mu_0$ и $\varepsilon_r = \varepsilon/\varepsilon_0$ – соответственно относительные магнитная и электрическая постоянные; $k_0 = \omega\sqrt{\mu_0\varepsilon_0}$ и $Z_0 = \sqrt{\mu_0/\varepsilon_0}$ – волновое число и волновое сопротивление (импеданс вакуума).

Используя уравнение (7) и граничные условия (5), (6), можно получить вариационную форму волнового уравнения [6, 8]:

$$\begin{aligned} \int_{\Omega} \left[\frac{1}{\mu_r} (\nabla \times \mathbf{W}_j) \cdot (\nabla \times \mathbf{E}) - k_0^2 \varepsilon_r \mathbf{W}_j \cdot \mathbf{E} \right] d\Omega = \int_{\Gamma_D} \frac{1}{\mu_r} (\mathbf{n} \times \mathbf{W}_j) (\nabla \times \mathbf{E}) d\Gamma - \\ - \int_{\Gamma_N} \left[\frac{ik_0}{\eta_r} (\mathbf{n} \times \mathbf{W}_j) (\nabla \times \mathbf{E}) + \mathbf{W}_j \cdot \mathbf{K}_N \right] d\Gamma - ik_0 Z_0 \int_{\Omega} \mathbf{W}_j \cdot \mathbf{J}_{\text{imp}} d\Omega, \end{aligned} \quad (8)$$

где \mathbf{W}_j – весовая функция.

2. Тетраэдральный конечный элемент для моделирования распределения электромагнитного поля в ближней зоне сферической наночастицы

Дискретизацию расчетной области Ω будем проводить конечными элементами в форме тетраэдра с тремя степенями свободы в каждом узле (E_x, E_y, E_z). Будем аппроксимировать значения искоемых функций линейными полиномами:

$$E_x(x, y, z) = \alpha_1 + \alpha_2 x + \alpha_3 y + \alpha_4 z, \quad E_y(x, y, z) = \alpha_5 + \alpha_6 x + \alpha_7 y + \alpha_8 z,$$

$$E_z(x, y, z) = \alpha_9 + \alpha_{10} x + \alpha_{11} y + \alpha_{12} z;$$

$$\mathbf{E} = [\mathbf{A}] \boldsymbol{\alpha}, \quad (9)$$

где

$$\mathbf{E}^T = \{E_x \quad E_y \quad E_z\}; \quad [\mathbf{A}] = \begin{bmatrix} 1 & x & y & z & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & x & y & z & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & x & y & z \end{bmatrix}; \quad \boldsymbol{\alpha}^T = \{\alpha_1 \quad \alpha_2 \quad \dots \quad \alpha_{12}\}.$$

Так как выражение (3) имеет место для любой точки тетраэдра, для его узлов получим

$$\mathbf{E}^e = [\mathbf{B}] \boldsymbol{\alpha}, \quad (10)$$

где

$$[\mathbf{B}] = \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \\ \mathbf{B}_3 \\ \mathbf{B}_4 \end{bmatrix}; \quad [\mathbf{B}_i] = \begin{bmatrix} 1 & x_i & y_i & z_i & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & x_i & y_i & z_i & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & x_i & y_i & z_i \end{bmatrix};$$

x_i, y_i, z_i ($i = \overline{1,4}$) – координаты узлов тетраэдра.

Полагая известным вектор \mathbf{E}^e , найдем из (10) вектор коэффициентов $\boldsymbol{\alpha}$:

$$\boldsymbol{\alpha} = [\mathbf{B}]^{-1} \mathbf{E}^e, \quad (11)$$

где

$$[\mathbf{B}]^{-1} = \frac{1}{6V} \begin{bmatrix} a_1 & 0 & 0 & a_2 & 0 & 0 & a_3 & 0 & 0 & a_4 & 0 & 0 \\ b_1 & 0 & 0 & b_2 & 0 & 0 & b_3 & 0 & 0 & b_4 & 0 & 0 \\ c_1 & 0 & 0 & c_2 & 0 & 0 & c_3 & 0 & 0 & c_4 & 0 & 0 \\ d_1 & 0 & 0 & d_2 & 0 & 0 & d_3 & 0 & 0 & d_4 & 0 & 0 \\ 0 & a_1 & 0 & 0 & a_2 & 0 & 0 & a_3 & 0 & 0 & a_4 & 0 \\ 0 & b_1 & 0 & 0 & b_2 & 0 & 0 & b_3 & 0 & 0 & b_4 & 0 \\ 0 & c_1 & 0 & 0 & c_2 & 0 & 0 & c_3 & 0 & 0 & c_4 & 0 \\ 0 & d_1 & 0 & 0 & d_2 & 0 & 0 & d_3 & 0 & 0 & d_4 & 0 \\ 0 & 0 & a_1 & 0 & 0 & a_2 & 0 & 0 & a_3 & 0 & 0 & a_4 \\ 0 & 0 & b_1 & 0 & 0 & b_2 & 0 & 0 & b_3 & 0 & 0 & b_4 \\ 0 & 0 & c_1 & 0 & 0 & c_2 & 0 & 0 & c_3 & 0 & 0 & c_4 \\ 0 & 0 & d_1 & 0 & 0 & d_2 & 0 & 0 & d_3 & 0 & 0 & d_4 \end{bmatrix};$$

V – объем элементарного тетраэдра;

$$a_i = (-1)^{i+1} \begin{vmatrix} x_j & y_j & z_j \\ x_k & y_k & z_k \\ x_n & y_n & z_n \end{vmatrix}; \quad b_i = (-1)^i \begin{vmatrix} 1 & y_j & z_j \\ 1 & y_k & z_k \\ 1 & y_n & z_n \end{vmatrix};$$

$$c_i = (-1)^{i+1} \begin{vmatrix} 1 & x_j & z_j \\ 1 & x_k & z_k \\ 1 & x_n & z_n \end{vmatrix}; \quad d_i = (-1)^i \begin{vmatrix} 1 & x_j & y_j \\ 1 & x_k & y_k \\ 1 & x_n & y_n \end{vmatrix};$$

i, j, k, n – номера вершин элементарного тетраэдра.

Поставив (11) в (9), получим

$$\mathbf{E} = [\mathbf{N}^e] \mathbf{E}^e, \quad (12)$$

где

$$[\mathbf{N}^e] = [\mathbf{A}][\mathbf{B}]^{-1} - \quad (13)$$

матрица, содержащая функции формы конкретного конечного элемента.

Электрическое поле \mathbf{E} в области Ω может быть найдено по формуле

$$\mathbf{E} = \sum_{l=1}^{N_{\text{узл}}} \mathbf{N}_l \mathbf{E}_l + \sum_{l=1}^{N_D} \mathbf{N}_l^D \mathbf{E}_l^D, \quad (14)$$

где $N_{\text{узл}}$ – число уникальных узлов в дискретизованной области, исключая те узлы, которые расположены на Γ_D ; \mathbf{E}_l – значение электрического поля в l -м узле; \mathbf{N}_l – функция формы для l -го узла; N_D – количество узлов на Γ_D ; \mathbf{N}_l^D и \mathbf{E}_l^D – соответственно векторная базисная функция и значение поля на Γ_D .

В качестве весовых функций \mathbf{W}_j в (8) примем \mathbf{N}_j , т. е.

$$\mathbf{W}_j = \mathbf{N}_j, \quad j = \overline{1, N_{\text{узл}}}. \quad (15)$$

Подставив (14) и (15) в (8), получим

$$\sum_{l=0}^{N_{\text{узл}}} \mathbf{K}_{jl} \mathbf{E}_l = \mathbf{b}_j, \quad j = \overline{1, N_{\text{узл}}}, \quad (16)$$

$$\text{где } \mathbf{K}_{jl} = \int_{\Omega} \left[\frac{1}{\mu_r} (\nabla \times \mathbf{N}_j) \cdot (\nabla \times \mathbf{N}_l) - k_0^2 \varepsilon_r \mathbf{N}_j \cdot \mathbf{N}_l \right] d\Omega + ik_0 \int_{\Gamma_N} \frac{1}{\eta_r} (\mathbf{n} \times \mathbf{N}_j) (\nabla \times \mathbf{N}_l) d\Gamma; \quad (17)$$

$$\mathbf{b}_j = -ik_0 Z_0 \int_{\Omega} \mathbf{N}_j \cdot \mathbf{J}_{\text{imp}} d\Omega - \int_{\Gamma_N} \mathbf{N}_j \cdot \mathbf{K}_N d\Gamma - \sum_{l=1}^{N_D} \mathbf{E}_l^D \int_{\Omega} \left[\frac{1}{\mu_r} (\nabla \times \mathbf{N}_j) \cdot (\nabla \times \mathbf{N}_l^D) - k_0^2 \varepsilon_r \mathbf{N}_j \cdot \mathbf{N}_l^D \right] d\Omega. \quad (18)$$

Система линейных алгебраических уравнений (16) представляет собой основное уравнение МКЭ. Функции формы в (17) для каждого конечного элемента вычисляются точно по формуле (13).

3. Учет граничных условий

Главная сложность конечно-элементного анализа для 3D-случая заключается в том, что необходимо решить, как бороться с окружающим безграничным открытым пространством вокруг исследуемого объекта (частицы). Общий подход состоит в том, чтобы создать некоторую ограничивающую поверхность [6]. Внутри такой искусственной области электрическое поле удовлетворяет волновому уравнению (7) при условии, что источник падающего поля находится вне ограничивающей поверхности. В дополнение к этому для однозначного решения задачи о распределении электромагнитных полей требуется определить граничное условие на ограничивающей поверхности. Данное граничное условие должно быть сформулировано таким образом, чтобы обеспечивалось распространение рассеянного поля без каких-либо искажений и отражений при переходе через ограничивающую поверхность, т. е. граничное условие должно быть «прозрачно» для рассеянного поля.

Воспользуемся излучением Зоммерфельда для определения граничного условия [4, 6]. Это позволяет принять, что если ограничивающая поверхность располагается достаточно далеко от рассеивателя (сферы или другого объекта, размещенного внутри ограниченной области), то рассеянное поле удовлетворяет соотношению

$$\mathbf{n} \times \left(\frac{1}{\mu_r} \nabla \times \mathbf{E}^{\text{sc}} \right) + \frac{ik_0}{\eta_r} \mathbf{n} \times (\mathbf{n} \times \mathbf{E}^{\text{sc}}) \approx 0, \quad (19)$$

где \mathbf{E}^{sc} – рассеянное поле, которое является разницей между полным и падающим полями:

$$\mathbf{E}^{\text{sc}} = \mathbf{E} - \mathbf{E}^{\text{inc}}. \quad (20)$$

В этом случае граничное условие (6) можно переписать в виде

$$\mathbf{n} \times \left(\frac{1}{\mu_r} \nabla \times \mathbf{E}^D \right) + \frac{ik_0}{\eta_r} \mathbf{n} \times (\mathbf{n} \times \mathbf{E}^D) = \mathbf{U}^{\text{inc}}, \quad (21)$$

$$\text{где } \mathbf{U}^{\text{inc}} = \mathbf{n} \times \left(\frac{1}{\mu_r} \nabla \times \mathbf{E}^{\text{inc}} \right) + \frac{ik_0}{\eta_r} \mathbf{n} \times (\mathbf{n} \times \mathbf{E}^{\text{inc}}). \quad (22)$$

4. Конечно-элементное моделирование распределения электромагнитного поля

Для исследования распределения электромагнитного поля, создаваемого плоской монохроматической волной, в ближней зоне сферической наночастицы предлагается следующий алгоритм конечно-элементного моделирования:

Шаг 1. На первом этапе решения задачи, исходя из исходных данных, вычисляем \mathbf{J}_{imp} и \mathbf{E}^{inc} .

Шаг 2. По формуле (22) вычисляем \mathbf{U}^{inc} и полагаем $\mathbf{K}_N = \mathbf{U}^{\text{inc}}$.

Шаг 3. Решая задачу (21), находим \mathbf{E}^D .

Шаг 4. Подставляя найденные значения \mathbf{E}^D и \mathbf{K}_N в (18), формируем основное уравнение МКЭ (16).

Шаг 5. Учитываем условия Дирихле (5), исключая из (16) известные значения \mathbf{E} , тем самым понижая порядок системы уравнений.

Шаг 6. Решая систему линейных алгебраических уравнений (16), находим искомые значения \mathbf{E} .

5. Верификация предложенной математической модели

Для верификации предложенной математической модели рассматривалась задача о распределении электрического поля в ближней зоне сферической наночастицы (см. рис. 1), для которой известно аналитическое решение, полученное Ми [1]. Для сравнения результатов вычислений воспользуемся относительными единицами:

$$\mathbf{E}^* = \frac{|\mathbf{E}|^2}{|\mathbf{E}_0|^2},$$

где \mathbf{E}_0 – векторная амплитуда падающей на наночастицу волны.

Наночастицы радиусом 40 нм находятся в воздухе ($\epsilon = 1$), их центры совпадают с центром декартовой системы координат. В качестве металлической наночастицы рассматривается серебряная (Ag) со значением диэлектрической проницаемости материала $\epsilon_{NP} = -3,5191 + 0,5798i$ (рис. 2 и 4). Для диэлектрической наночастицы выбрано значение $\epsilon_{NP} = 3,5191 + 0,5798i$ (рис. 3 и 5). Для расчета параметров частиц использовалась модель Лоренца – Друде. Расчетная область дискретизировалась равновеликими тетраэдрами для каждой из зон (см. рис. 1).

Решение, полученное с помощью предложенного алгоритма, сравнивалось с решением на основе теории Ми (в областях вне наночастиц) [1]. Матрица жесткости вычислялась по формуле (17). Граничные условия находились из решения уравнений (21). Максимальная погрешность решений не превышала 10 %.

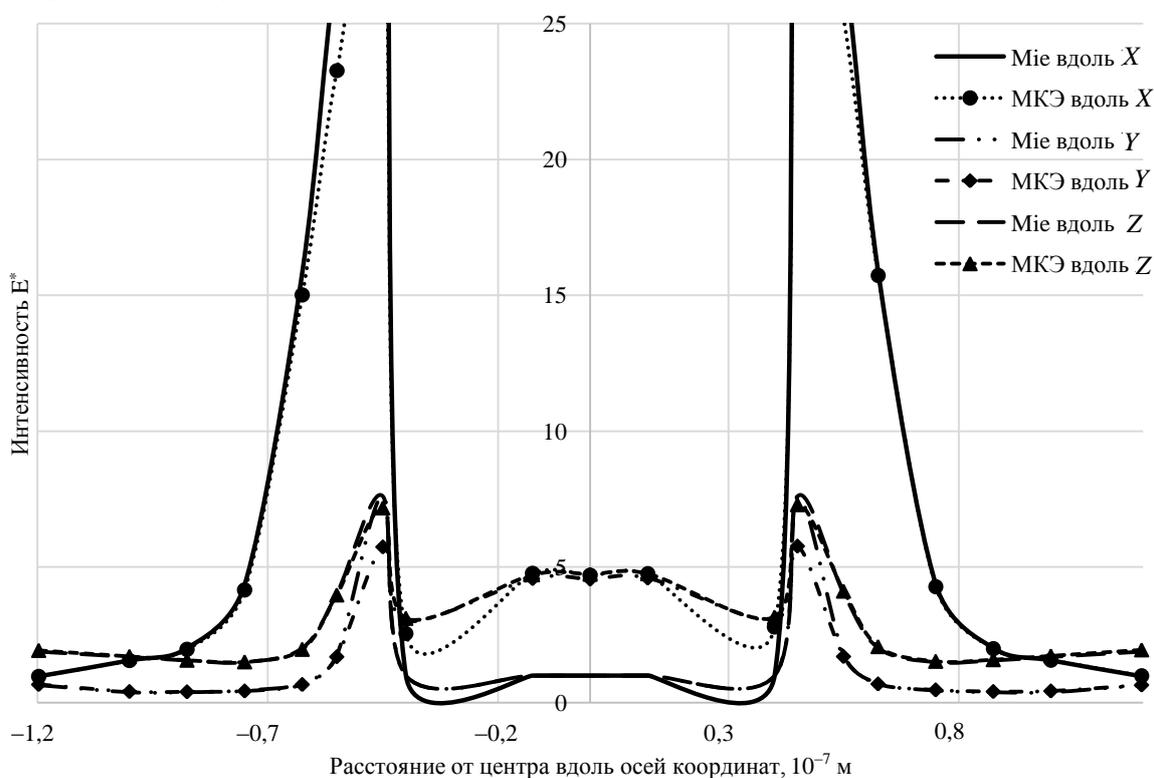


Рис. 2. Зависимость интенсивности электромагнитного поля в ближней зоне от расстояния от центра металлической наночастицы

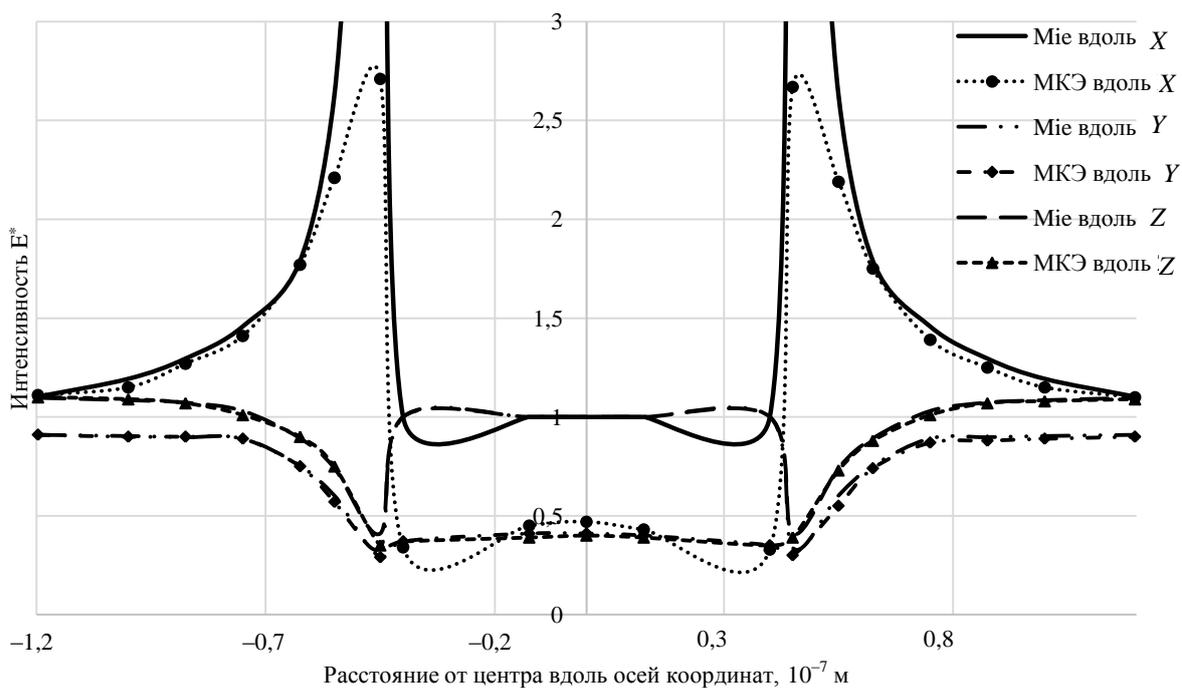


Рис. 3. Зависимость интенсивности электромагнитного поля в ближней зоне от расстояния от центра диэлектрической наночастицы

На рис. 2–5 различные кривые соответствуют изменению одной из координат при нулевых значениях двух других координат, т. е. рассматривается распределение поля в центральных сечениях сферической наночастицы, параллельных координатным осям. Возбуждение осуществлялось плоской линейно поляризованной вдоль оси X волной длиной 405 нм.

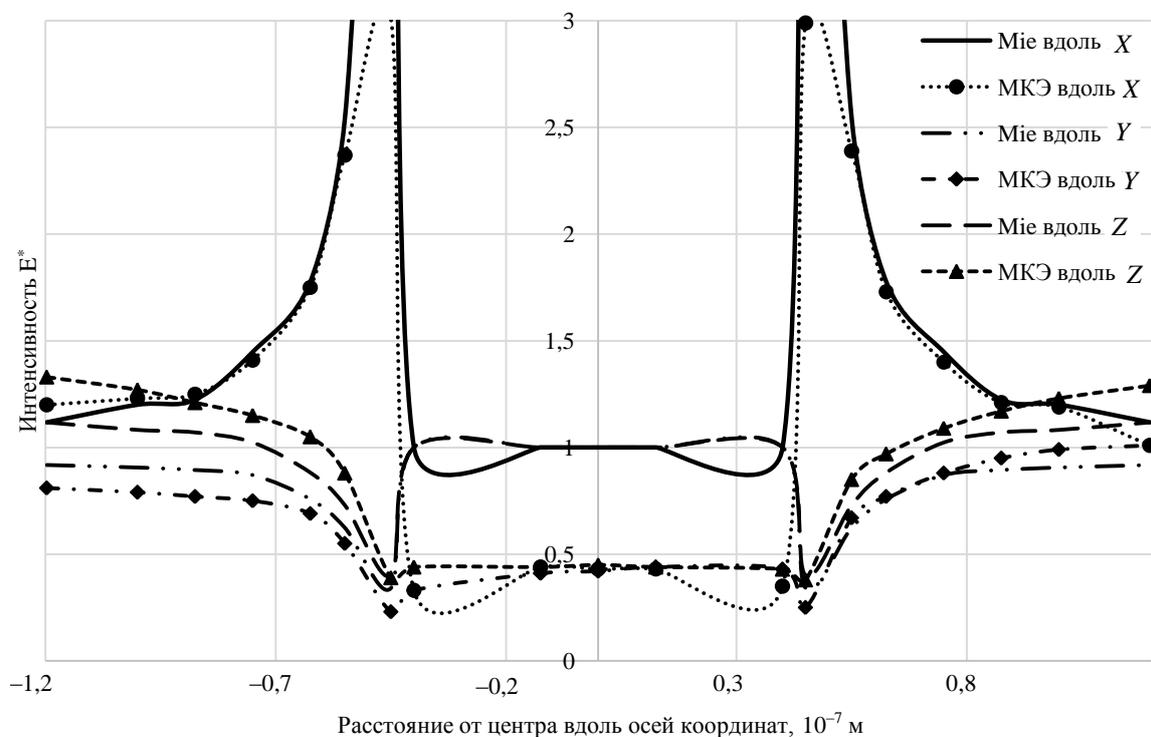


Рис. 4. Зависимость интенсивности электромагнитного поля в ближней зоне от расстояния от центра диэлектрической наночастицы в золь-гель-стекле

6. Вычислительный эксперимент

На основе созданной математической модели было разработано программное обеспечение (свидетельство о регистрации компьютерной программы ElectroMagnetic Field FEM Modeler (EMFFM) в Национальном центре интеллектуальной собственности № 514, Республика Беларусь, 2013), с помощью которого проведено исследование распределения электромагнитного поля в ближней зоне сферических наночастиц в золь-гель-стекле (SiO_2) [2].

Характеристики серебряных наночастиц, материалов вне наночастиц и условий возбуждения для данных на рис. 4 и 5 выбирались в областях значений параметров, реализуемых для нанокompозитов на основе высококремнеземных золь-гель-стекол [2, 7].

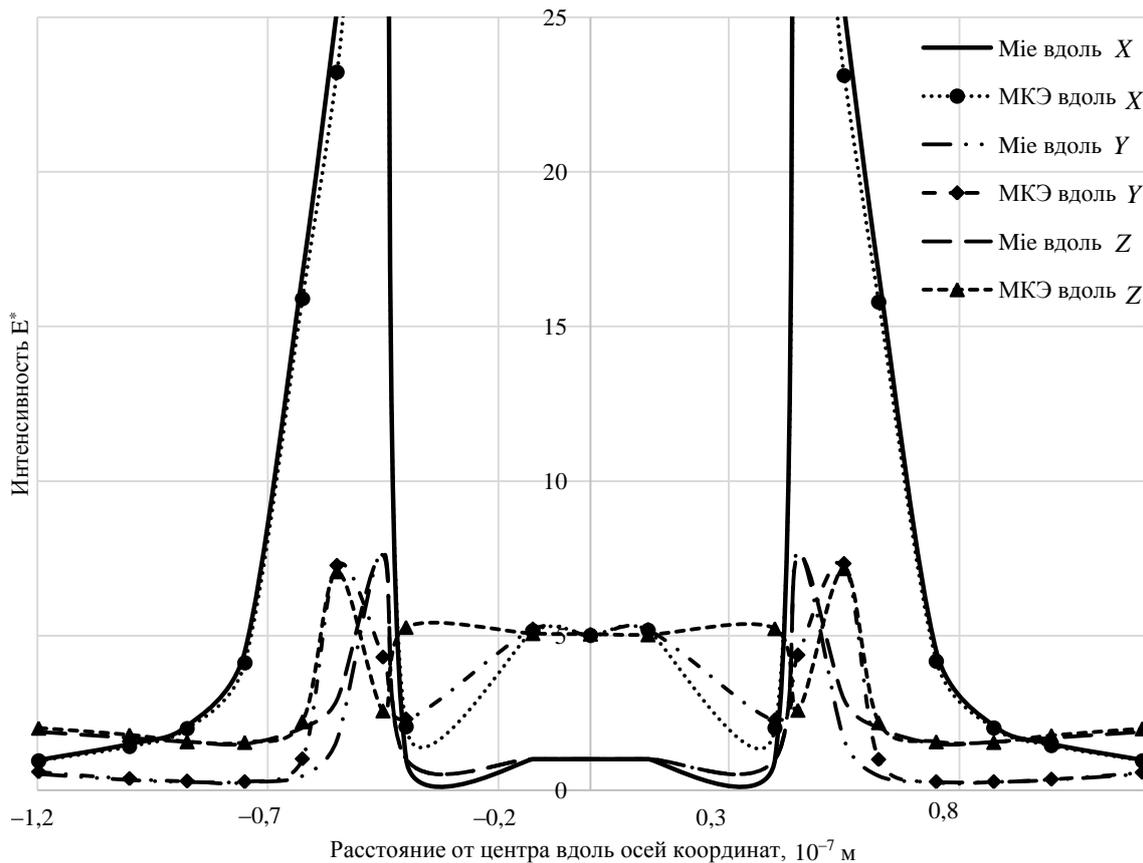


Рис. 5. Зависимость интенсивности электромагнитного поля в ближней зоне от расстояния от центра металлической наночастицы в золь-гель-стекле

Заключение

Предлагаемая математическая модель и численный алгоритм ее реализации могут быть использованы для исследования распределения электромагнитного поля в ближней зоне сферической наночастицы.

Согласно проведенному вычислительному эксперименту (рис. 2–5) при удалении от поверхности наночастицы на расстояние порядка радиуса в направлении вдоль одной из координатных осей относительная амплитуда электрического поля быстро уменьшается (увеличивается) для параллельных (перпендикулярных) полю E_0 направлений. Усиление ближнего поля в окружающей наночастицу среде может иметь место как для металлической (см. рис. 2), так и для диэлектрической (см. рис. 3) наночастиц (в основном для направлений, параллельных плоскости поляризации возбуждающего излучения). При этом максимальные значения величины E^* для серебряной наночастицы на порядок и более превосходят аналогичные значения для диэлектрической наночастицы.

Получены соотношения для аналитического вычисления локальных матриц жесткости по формуле (18), что минимизирует погрешность при нахождении глобальной матрицы жесткости и позволяет ускорить процесс ее формирования.

Достоинством предлагаемой математической модели и методики ее применения является использование узловых конечных элементов, что значительно упрощает процесс формирования глобальной матрицы жесткости в сравнении с применением элементов реберных типов. Это достигается за счет отсутствия необходимости учета порядка и ориентации обхода ребер конечного элемента в пространстве.

Список литературы

1. Хюлст, Г. ван де. Рассеяние света малыми частицами / Г. ван де Хюлст. – М. : Изд-во иностр. лит., 1961. – 537 с.
2. Алексеенко, А.А. Функциональные материалы на основе диоксида кремния, получаемые золь-гель-методом / А.А. Алексеенко, А.А. Бойко, Е.Н. Подденежный. – Гомель : ГГТУ им. П.О. Сухого, 2008. – 183 с.
3. Наноструктурные материалы – 2012: Россия – Украина – Беларусь / Тез. докл. III Междунар. науч. конф., СПб., 19–22 нояб. 2012 г. – СПб. : Лема, 2012.
4. Климов, В.В. Наноплазмоника / В.В. Климов. – М. : Физматлит, 2009. – 480 с.
5. Comparison of numerical methods in near-field computation for metallic nanoparticles / M. Karamehmedovic [et al.] // *Optic Express*. – 2011. – Vol. 19, no. 9. – P. 8939–8953.
6. Jianming, J. Theory and Computation of Electromagnetic Fields / J. Jianming. – John Wiley & Sons, 2010. – 616 p.
7. Курочка, К.С. Построение программно-аппаратного комплекса для исследования распределения электромагнитного поля в ближней зоне диэлектрических и металлических наночастиц / К.С. Курочка, О.Д. Асенчик, Е.Г. Стародубцев // Докл. БГУИР. – 2013. – № 3(81). – С. 64–70.
8. Jianming, J. The Finite Element Method in Electromagnetics / J. Jianming. – 2nd ed. – N. Y. : Wiley, 2002. – 780 p.
9. Курочка, К.С. Моделирование распространения электромагнитного поля в ближней зоне металлической частицы методом конечных элементов / К.С. Курочка, А.А. Кухаренко // Материалы VII науч. конф. «Математическое и имитационное моделирование систем (МОДС – 2012)», 25–28 июня 2012 г., Украина / Черниговский гос. техн. ун-т. – Чернигов – Жукин, 2012. – С. 36–39.
10. Sadiku, M.N.O. Numerical Techniques in Electromagnetics / M.N.O. Sadiku. – 2nd ed. – London : CRC Press, 2003.

Поступила 27.08.2016

*Гомельский государственный технический
университет им. П.О. Сухого,
пр. Октября, 48
e-mail: kurochka@gstu.by*

K.S. Kurachka

FINITE-ELEMENT MODELING OF ELECTROMAGNETIC FIELD DISTRIBUTION IN THE NEAREST ZONE OF A SPHERICAL NANOPARTICLE

Some problems of the finite element modeling of electromagnetic field distribution in the nearest zone of spherical metallic and dielectric nanoparticles caused by a plane wave are considered. The appropriate mathematical models, algorithms and software were developed, its verification is carried out. The discrepancy between the results of research of the mathematical model and the existing results of calculations using analytical formulas does not exceed 10 %.

ОБРАБОТКА СИГНАЛОВ, ИЗОБРАЖЕНИЙ И РЕЧИ

УДК 343.98

А.А. Борискевич

ВНЕДРЕНИЕ ИДЕНТИФИКАЦИОННЫХ ДАННЫХ В ИЗОБРАЖЕНИЯ
НА ОСНОВЕ СИНУСОИДАЛЬНЫХ РЕШЕТОК
ДЛЯ МОБИЛЬНЫХ ПРИЛОЖЕНИЙ

Предлагается метод оптически визуализируемого блочно-структурного маркирования изображений, основанный на моделях внедряемого сообщения и маркированного изображения и робастном текстурно-корреляционном извлечении идентификационной информации. Представляются результаты компьютерного моделирования.

Введение

Широкое распространение мобильных телефонов с камерой привело к появлению множества двумерных штриховых кодов (AztecCode, DataMatrix, PDF-417, QR-кодов) [1–4]. Пользователь, имеющий мобильное устройство, может быстро извлечь информацию, сохраненную в двумерном штрихкоде, такую как контактная информация, реклама или URL (стандартизированный способ записи адреса ресурса в сети Интернет). Однако двумерные штрихкоды лишены визуальной информации и состоят из белых и черных блоков, что затрудняет их использование в полиграфических продуктах и цифровых печатных технологиях, например, с целью усложнения подделки (фальсификации) ID-карт, контролирующей доступ к конфиденциальной информации, и внедрения идентификационных данных в изображение для прямой связи его с соответствующим контентом в Интернете [1, 4]. Цифровое маркирование для мобильных медиатехнологий должно удовлетворять следующим требованиям: вложенная информация не должна искажать визуальные характеристики изображения, она должна быть перцептуально невидимой и устойчивой к условиям печати и считывания [5–9], т. е. устойчивой к цифроаналоговым и аналого-цифровым воздействиям.

Для устранения указанных недостатков и обеспечения контроля доступа к информационным ресурсам с использованием мобильных цифровых медиатехнологий предложен метод блочно-структурного маркирования и считывания маркированных изображений с помощью мобильных устройств (рис. 1). Метод основан на генерации блоков эталонных синусоидальных решеток в качестве оптически визуализируемых образов с тремя идентификационными параметрами (ориентацией, периодом и размером) для кодирования информационного сообщения, гибридном правиле внедрения синусоидальных блоков, формировании рамки обнаружения маркированного изображения для оптического захвата изображения и текстурной обработки маркированного изображения для корреляционного декодирования сообщения с использованием эталонных синусоидальных блоков.

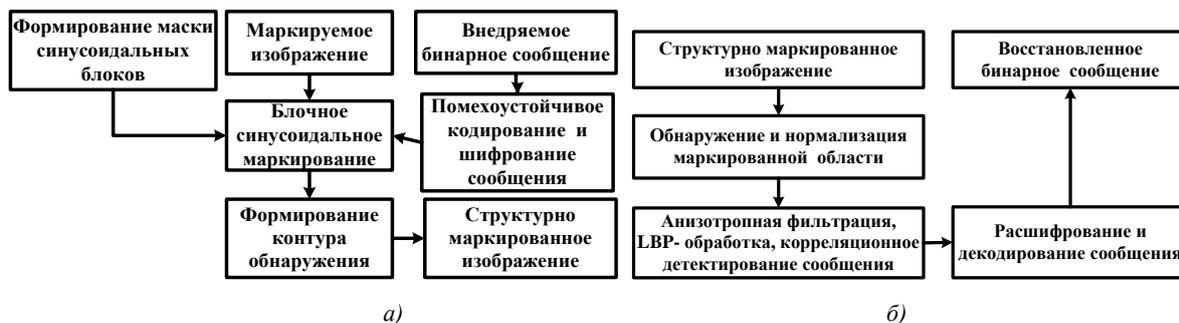


Рис. 1. Блок-схема метода блочно-структурного маркирования и считывания маркированных изображений с помощью мобильных устройств: а) синусоидальное маркирование изображений; б) текстурно-корреляционное извлечение идентификационной информации

1. Метод оптически визуализируемого блочного маркирования изображений

Внедряемое секретное сообщение с K битами информации имеет вид $b = b_0, \dots, b_k, \dots, b_{K-1}$, где $b_k \in \{0, 1\}$. Для предотвращения битовых ошибок формирование избыточной знакопеременной двоичной последовательности секретного сообщения $b_R = ((b_0, \dots, b_0)_L \dots (b_k, \dots, b_k)_L \times (b_{K-1}, \dots, b_{K-1})_L)$ осуществляется с помощью кода с повторением $(L, 1)$ при соблюдении условия $LK \leq C$, где C – максимальное количество битов, которое может быть внедрено в изображение. Использование операции повторения снижает в L раз объем внедряемой информации в маркируемое изображение. Данный код может корректировать $\lfloor (L-1)/2 \rfloor$ ошибок в кодовом слове длиной L . Декодирование кода с повторением осуществляется с помощью мажоритарного правила, основанного на выборе значений битов из кодовой последовательности, которые встречаются наиболее часто. Сообщение b_R шифруется с помощью блочных и поточных алгоритмов шифрования (AES (Advanced Encryption Standard), RC4) [10–13] или генераторов детерминированного хаоса [14] с длиной секретного ключа не менее 128 бит.

Закодированное и зашифрованное бинарное сообщение заменяется последовательностью синусоидальных блоков с идентификационными параметрами (ориентацией θ , частотой $f_{u,v}$ и размером $U \times V$), формируемых с помощью соотношения

$$S_{\theta\delta} = (S(u, v, \delta_1, \delta_2, f_u, f_v)_{U \times V} = \sin 2\pi(\delta_1 f_u u + \delta_2 f_v v) \mid u = 0, M/W - 1, v = 0, N/W - 1), \quad (1)$$

где δ_1 и δ_2 – параметры, определяющие значение угла ориентации $\theta = \arctg(\delta_2 f_v / \delta_1 f_u)$ синусоидальных решеток; $f_{u,v}$ и W – значение частоты решетки с линейными размерами $U = M/W$ и $V = N/W$, обратно пропорциональное ее периоду в пикселах, и количество блоков, на которые разбивается маркируемое изображение, соответственно.

Для формирования двух и четырех синусоидальных решеток параметры δ_1 и δ_2 принимают следующие значения: $\delta_1 = 1$, $\delta_2 \in \{-1, 0; 1, 0\}$ и $\delta_1 = 1$, $\delta_2 \in \{1, 0; -1, 0; 0, 5; -0, 5\}$ (рис. 2) соответственно. Для оптически визуализируемого представления пар двоичных символов (00, 01, 10 и 11) используются решетки с четырьмя угловыми ориентациями ($+45^\circ$, -45° , $+25^\circ$ и -25°) (рис. 2).

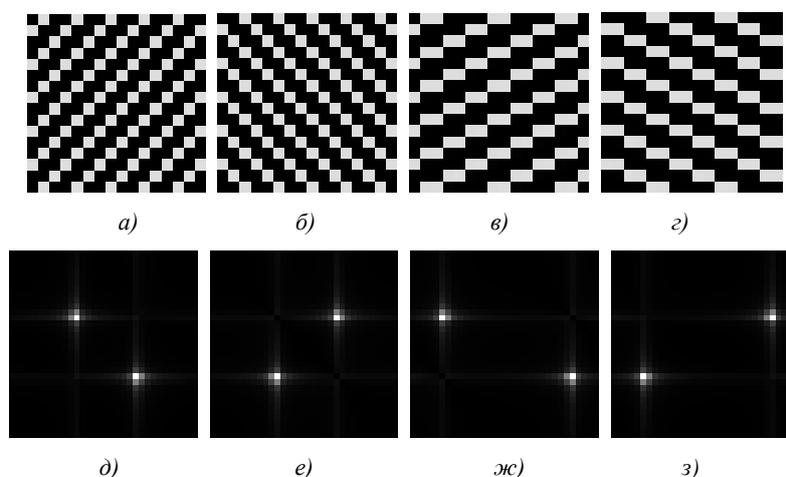


Рис. 2. Пространственно-частотные особенности эталонных синусоидальных блоков размером 16×16 :

а) $S_{0,1}$; б) $S_{0,-1}$; в) $S_{0,0,5}$; г) $S_{0,-0,5}$; д) – з) амплитудные Фурье-спектры решеток

Маркирование изображения $I = (I(m, n))$ осуществляется посредством гибридного правила внедрения секретного сообщения в виде маркирующей маски $S = (S(m, n))$, состоящей из синусоидальных блоков размером $U \times V$ с различными угловыми ориентациями:

$$I_{w_{\alpha\beta}}(m, n) = I(m, n) + \alpha S(m, n) \pm \beta I(m, n)S(m, n), \quad (2)$$

где $I_{w_{\alpha\beta}}(m, n)$ – значения пикселей маркированного изображения $I_{w_{\alpha\beta}}$; α и β – коэффициенты, характеризующие силу внедрения и контролирующие изменения значений пикселей, вызываемых соответственно аддитивной $\alpha S(m, n)$ и мультипликативной $\beta I(m, n)S(m, n)$ составляющими.

Из выражения (2) следует, что аддитивная, мультипликативная и обобщенная схемы вложения реализуются при $(\beta = 0, \alpha \neq 0)$, $(\beta \neq 0, \alpha = 0)$ и $(\beta \neq 0, \alpha \neq 0)$ соответственно. Возможность изменения знака перед β позволяет минимизировать величину деградации маркированного изображения не только для малых, но и для больших значений параметров α и β и увеличить энергию внедряемого сообщения.

Для оптически визуализируемого представления отдельных двоичных символов (0 и 1) используются решетки с двумя угловыми ориентациями ($+45^\circ$ и -45°). На рис. 3 показан процесс маркирования изображения Ф. Скорины синусоидальными блоками с двумя угловыми ориентациями ($+45^\circ$ и -45°).

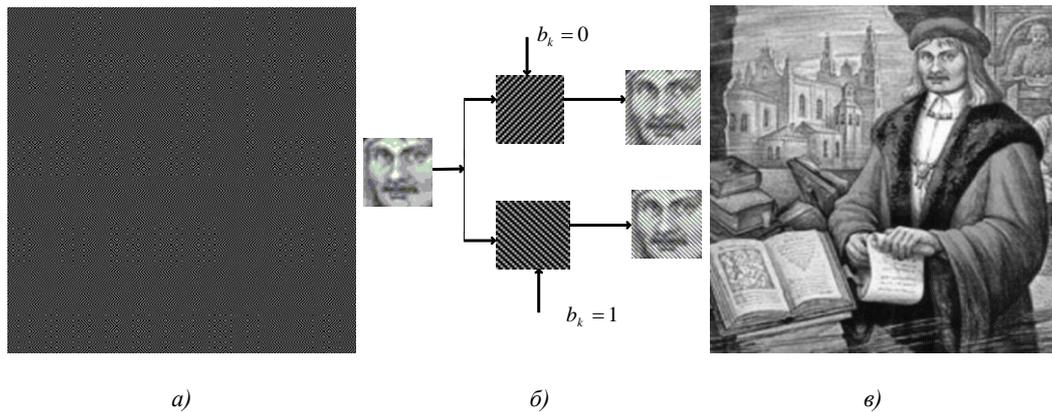


Рис. 3. Синусоидальное блочное маркирование изображения Ф. Скорины:
а) маркирующая маска кодирования, состоящая из синусоидальных блоков различной угловой ориентации; б) маркирование фрагмента изображения; в) маркированное изображение

С целью обнаружения маркированного изображения формируется прерывистая рамка определенной толщины и размера $M_{dc} \times N_{dc}$. Линии рамки изображения состоят из белого и черного слоев, каждый из которых содержит 3 пиксела. Маркированное изображение с рамкой обнаружения имеет вид $\tilde{I}_{dc} = (\tilde{I}_{dc}(m, n) | m = \overline{0, M_{dc} - 1}, n = \overline{0, N_{dc} - 1})$, где $M_{dc} = M + \Delta M_{dc}$, $N_{dc} = N + \Delta N_{dc}$, $\Delta M_{dc} = \Delta N_{dc} = 12$ пикселей.

Для обеспечения устойчивости к условиям печати и оптического считывания предложена робастная процедура извлечения маркирующей информации мобильными устройствами, основанная на робастном обнаружении рамки, нормализации маркированного изображения, предварительной обработке нормализованного изображения с использованием робастного ЛВР-признака изображения, итеративной анизотропной фильтрации и декодировании сообщения по выделенным синусоидальным блокам.

Обнаружение и нормализация маркированного изображения $I_{dc} = (I_{dc}(m, n) | m = \overline{0, M_{dc} - 1}, n = \overline{0, N_{dc} - 1})$ с детектируемой рамкой основаны на трех процедурах: формировании бинарного контурного изображения, локализации граничных точек изображения и компенсации геометрических искажений изображения.

Для обработки низкоконтрастных маркированных изображений с целью достижения инвариантности к условиям печати и съемки используется процедура формирования бинарного контурного изображения, основанная на методе Оцу (Otsu) [15] и алгоритме обнаружения границ Канны (Canny) [16].

Процесс локализации граничных точек маркированного изображения базируется на представлении в параметрическом пространстве $\rho\theta$ бинарного контурного изображения $\tilde{B}_H = \tilde{B}_{SG\theta} = (\tilde{B}_{SG}(\rho_i, \theta_i))$, где ρ_i – расстояние между прямой, проходящей через точку $\tilde{B}_{GSxy}(i, j)$ бинарного изображения и началом координат, θ_i – угол между перпендикуляром к линии и вертикальной осью X; на преобразовании Хафа [17] и селективной коррекции позиций граничных точек [18]. Данная коррекция основана на выборе четырех угловых граничных точек из граничных линий изображения, определении условия проверки окрестности 3×3 выбранных граничных точек на контурном бинарном изображении \tilde{B}_{GSxy} и поиске в восьми направлениях двух кандидатов на граничную точку посредством вычисления весовых коэффициентов.

Процедура компенсации геометрических искажений маркированного изображения основана на использовании обнаруженных четырех угловых граничных точек, формирующих вектор бесконечно удаленной линии, которая проходит через исчезающие точки, и формировании матрицы проективного преобразования маркированного изображения [19] в виде иерархической цепочки трех преобразований (изометрии и подобия, аффинного и проекционного), корректирующей координаты пикселей исходного изображения в нормализованные координаты.

Предварительная обработка нормализованного маркированного изображения заключается в удалении вертикальных и горизонтальных линий на основе линейной пространственной фильтрации и формировании LBP-изображения [20, 21]. Предложенный метод с размером примитива 3×3 для вычисления быстрого инвариантного к повороту LBP-признака позволяет получить выигрыш в скорости вычисления на уровне 66–72 % по отношению к методу без ускорения за счет сокращения относительного количества операций сравнения в два раза и относительного количества пикселей, к которым осуществляется доступ при вычислении LBP-признака, в 4,3 раза [21].

Анизотропная фильтрация нормализованного маркированного изображения I_n осуществляется с помощью пространственного фильтра с маской весовых коэффициентов

$$H = \begin{pmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{pmatrix} \text{ с целью удаления вертикальных и горизонтальных линий } I_n. \text{ Усиление диа-}$$

гональных линий маркированного изображения I_n с удаленными вертикальными и горизонтальными линиями осуществляется с помощью линейной пространственной фильтрации синусоидального блока маркированного изображения с помощью двух масок-фильтров

$$H_1 = \begin{pmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 2 & 0 \\ -1 & 0 & 2 & 0 & -1 \\ 0 & 2 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{pmatrix} \text{ и } H_2 = \begin{pmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & 2 & 0 & -1 & 0 \\ -1 & 0 & 2 & 0 & -1 \\ 0 & -1 & 0 & 2 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{pmatrix}, \text{ сравнения энергии } E_{-d} \text{ и } E_d$$

двух блоков, отфильтрованных масками-фильтрами H_1 и H_2 , с помощью соотношения

$$E = \begin{cases} E_d & \text{при } E_d > E_{-d}, \\ E_{-d} & \text{при } E_d < E_{-d}, \end{cases}$$

где $E_{-d} = \sum_{u=0}^{U-1} \sum_{v=0}^{V-1} |e_{-d}|^2$ и $E_d = \sum_{u=0}^{U-1} \sum_{v=0}^{V-1} |e_d|^2$ – значения энергии матрицы отфильтрованных блоков e_{-d} и e_d с ориентациями -1 и $+1$ соответственно, и выбора из двух блоков блока с большей энергией.

Для повышения точности детектирования синусоидальных решеток до 100 % и устойчивости к условиям печати и съемки осуществляется вычисление LBP-изображения для маркированного изображения с усиленными диагональными линиями [21, 22].

Значения бит информационного сообщения определяются с помощью вычисления коэф-

$$f(S_{\theta\delta}, I_{\theta\delta}) = \frac{\sum_{u=0}^{U-1} \sum_{v=0}^{V-1} (S_{\theta\delta}(u, v) - E(S_{\theta\delta})) (I_{\theta\delta}(u, v) - E(I_{\theta\delta}))}{\sqrt{\sum_{u=0}^{U-1} \sum_{v=0}^{V-1} (S_{\theta\delta}(u, v) - E(S_{\theta\delta}))^2 \sum_{u=0}^{U-1} \sum_{v=0}^{V-1} (I_{\theta\delta}(u, v) - E(I_{\theta\delta}))^2}} \quad \text{меж-}$$

ду эталонными $S_{\theta\delta} = (S_{\theta\delta}(u, v))$ и внедренными $I_{\theta\delta} = (I_{\theta\delta}(u, v))$ синусоидальными блоками маркированного изображения и выбора пары битов, соответствующих максимальному значению коэффициента корреляции

$$b_k b_{k+1} = \begin{cases} 10 & \text{при } \max_{\theta\delta} f(S_{\theta\delta}, I_{\theta,1}), \\ 00 & \text{при } \max_{\theta\delta} f(S_{\theta\delta}, I_{\theta,-1}), \\ 11 & \text{при } \max_{\theta\delta} f(S_{\theta\delta}, I_{\theta,0.5}), \\ 01 & \text{при } \max_{\theta\delta} f(S_{\theta\delta}, I_{\theta,-0.5}), \end{cases} \quad (3)$$

где $E(S_{\theta\delta})$ и $E(I_{\theta\delta})$ – средние арифметические значения эталонных $S_{\theta\delta}$ и внедренных $I_{\theta\delta}$ синусоидальных блоков.

Избыточная битовая последовательность расшифровывается с использованием секретного ключа пользователя. Коррекция ошибок осуществляется с помощью соотношения

$$q_k = \text{round} \left(\sum_{l=0}^{L-1} b_{kl} / L \right), \quad \text{где } b_{kl} \text{ – } l\text{-й бит расширенного } k\text{-го сегмента внедряемого сообщения}$$

$$b_R = ((b_0, \dots, b_0)_L \dots (b_k, \dots, b_k)_L \dots (b_{K-1}, \dots, b_{K-1})_L).$$

2. Результаты моделирования

Компьютерный эксперимент проводился при следующих условиях: цифровые изображения размером 512×512 пикселей, 16-символьная текстовая строка «Boriskevich A.A.» в качестве секретного сообщения, каждый символ кодируется 8 битами в соответствии с форматом ASCII, размер синусоидального блока 16×16 , аддитивное, мультипликативное и обобщенное правила внедрения, камера мобильного телефона 3 Мп с разрешением 2048×1536 , печать с разрешающей способностью 600 и 1200 dpi, физический размер маркированного изображения 7×7 см. Емкость $C = (MN/UV)K_\delta$ внедрения маркируемого изображения размером $M \times N$ зависит от размера $U \times V$ синусоидальной решетки и количества битов K_δ внедряемого сообщения, представляемых решеткой с определенной угловой ориентацией.

На рис. 4–6 и в таблице представлены результаты моделирования разработанного метода оптически визуализируемого блочного маркирования изображений на основе синусоидальных решеток с идентификационными параметрами.

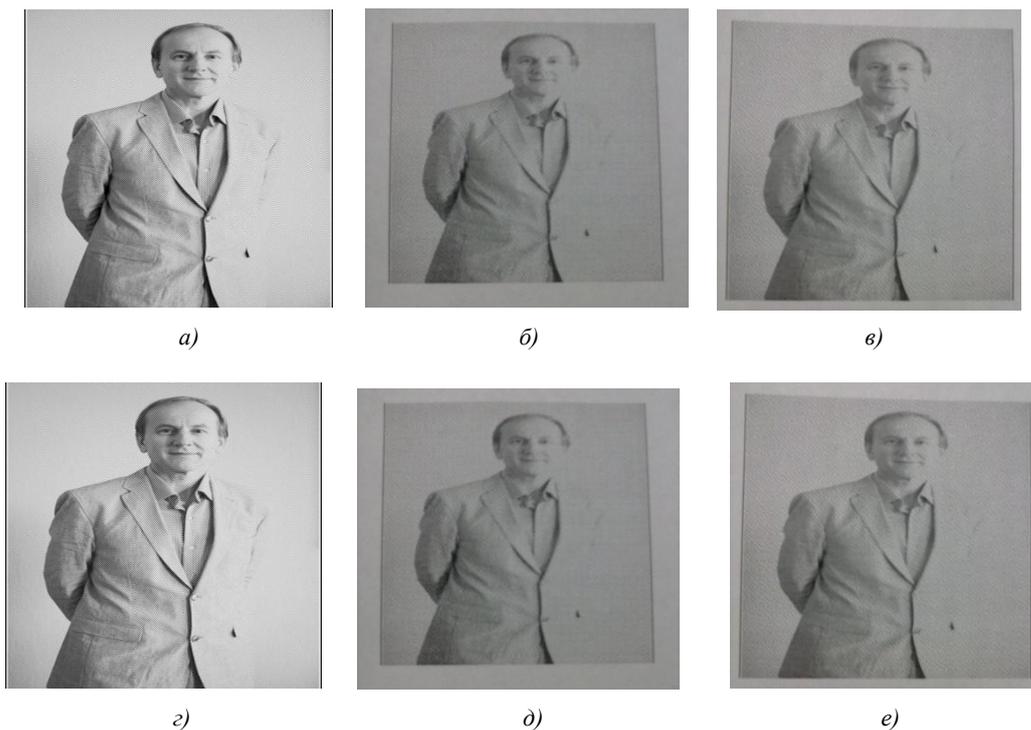


Рис. 4. Изображения размером 512×512 с различными правилами внедрения и качеством печати:
 а), з) цифровые изображения с мультипликативным и гибридным правилами внедрения;
 б), д) печатные изображения с мультипликативным и гибридным правилами внедрения при 600 dpi;
 в), е) печатные изображения с мультипликативным и гибридным правилами внедрения при 1200 dpi

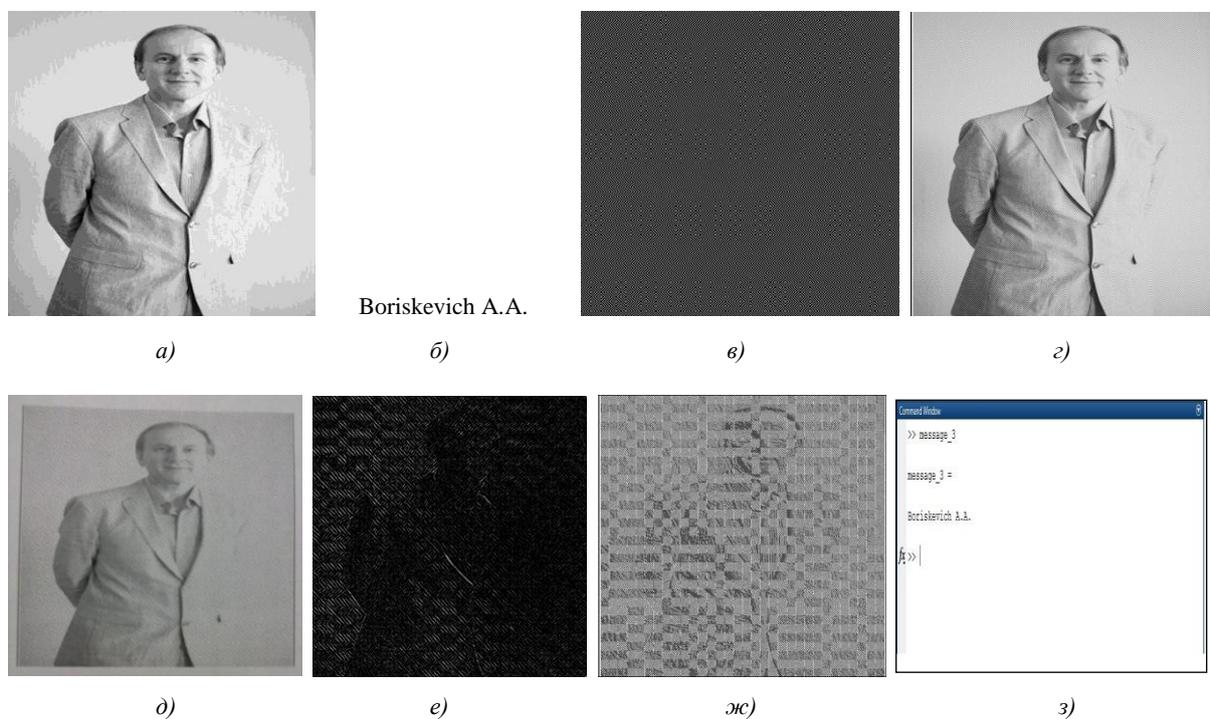


Рис. 5. Особенности блочного маркирования, считывания, обработки и декодирования изображения:
 а) цифровое маркируемое изображение; б) секретное сообщение; в) маркирующая маска синусоидальных блоков;
 з) цифровое маркированное изображение с детектируемой рамкой; д) печатное нормализованное маркированное изображение; е), ж) нормализованные маркированные изображения после анизотропной фильтрации и LBP-обработки; з) извлеченное сообщение

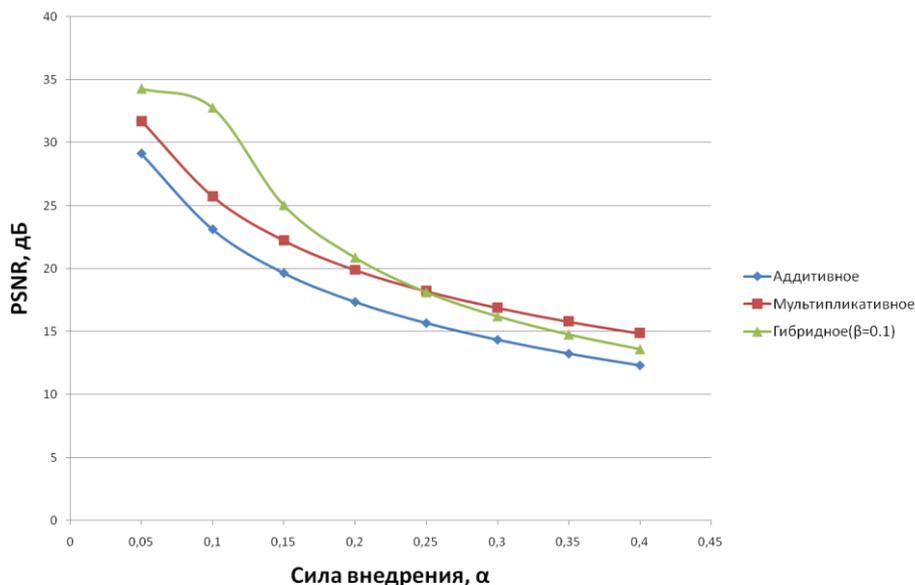


Рис. 6. Зависимость качества маркированного изображения в понятиях PSNR (пиковое отношение сигнал-шум) от силы внедрения α при фиксированном $\beta=0,1$ для аддитивного, мультипликативного и гибридного правил внедрения

Зависимость битовых ошибок от правила внедрения и качества печати

Правило внедрения	Количество точек на дюйм, dpi	Количество экспериментов	Битовые ошибки, %
Аддитивное	600	25	10,41
	1200	25	6,12
Мультипликативное	600	25	9,55
	1200	25	3,52
Гибридное	600	25	7,78
	1200	25	3,22

Для уменьшения процента ошибок до нуля при декодировании маркированного изображения целесообразно использовать анизотропную фильтрацию с определенным количеством повторений. Определено, что достаточно трех повторений анизотропной фильтрации.

Разработанный метод обеспечивает инвариантность к условиям печати и съемки (к повороту изображения в диапазоне $(-45^\circ, \dots, +45^\circ)$ и аддитивному изменению интенсивности изображения) и высокое качество маркированного изображения с $PSNR \in [30-34]$ дБ за счет предварительной обработки маркированного изображения и двухэтапной процедуры определения граничных угловых точек. Первый этап включает в себя формирование контурного бинарного изображения и обнаружение первичных позиций граничных точек, а второй этап – уточнение позиций граничных точек в локальной окрестности определенного размера вдоль граничных линий рамки и компенсацию геометрических искажений с помощью проективного преобразования при использовании гибридного правила внедрения с $\alpha \in [0,08-0,12]$ и $\beta = 0,1$.

Заключение

Разработан метод оптически визуализируемого блочного маркирования изображений и считывания маркированных изображений с помощью мобильных устройств. Отличительной особенностью метода является использование квадратных синусоидальных блоков в качестве кодовых образов двоичных символов внедряемого бинарного сообщения, модели маркированного изображения в виде алгебраической суммы аддитивной и мультипликативной яркостных составляющих и текстурно-корреляционного декодирования внедренных синусоидальных блоков, что обеспечивает высокое качество маркированного изображения ($PSNR > 30$ дБ), емкость

внедрения не менее 2^{10} байт, инвариантность к условиям сканирования (угловому положению (-45° , ..., $+45^\circ$), аддитивному изменению интенсивности изображения) и высокую точность декодирования синусоидальных решеток (100 %).

К установленным ограничениям метода относятся визуальное качество маркированного изображения (не менее 35 дБ); минимальный размер решетки (16×16 пикселей), минимальный период решетки (2 пиксела на период), максимальное количество угловых ориентаций (четыре наклонные решетки с углами наклона -45° , $+45^\circ$, -25° , $+25^\circ$).

Предложенный метод позволяет формировать маркированные изображения, несущие как визуальную, так и скрытую информацию, считываемую с экранов мониторов и с печатной продукции при помощи мобильных устройств.

Список литературы

1. Design of a Simple User Authentication Scheme Using QR Code for Mobile Device / Y. Lee [et al.] // Information Technology Convergence, Secure and Trust Computing, and Data Management. – 2012. – Vol. 180. – P. 241–248.
2. Ohbuchi, E. Barcode Readers Using the Camera Device in Mobile Phones / E. Ohbuchi, H. Hanaizumi, L.A. Hock // Proc. of the 2004 Intern. Conf. on Cyberworlds (CW'04). – Tokyo, Japan, 2004. – P. 260–265.
3. Gu, Y. QR Code Recognition Based on Image Processing / Y. Gu, W. Zhang // Intern. Conf. on Information Science and Technology (ICIST). – Jiangsu, China, 2011. – P. 733–736.
4. Liao, K. A Novel User Authentication Scheme Based on QR-Code / K. Liao, W. Lee // Journal of Networks. – 2010. – Vol. 5, no. 8. – P. 937–941.
5. Keskinarkaus, A. Digital watermarking techniques for printed images / A. Keskinarkaus. – Finland, University of Oulu, 2012. – 101 p.
6. Pramila, A. Toward an interactive poster using digital watermarking and a mobile phone camera / A. Pramila, A. Keskinarkaus, T. Seppanen // Journal of Signal, Image and Video processing. – 2012. – Vol. 6, no. 2. – P. 211–222.
7. Alattar, A.M. Smart Images using Digimarc's Watermarking Technology / A.M. Alattar // Proc. SPIE: Security and Watermarking of Multimedia Contents. – 2000. – Vol. 3971. – P. 264–273.
8. He, D. A Practical Print-scan Resilient Watermarking Scheme / D. He, Q. Sun // Proc. IEEE Intern. Conf. on Image Processing (ICIP). – Genova, Italy, 2005. – P. 257–260.
9. Image Fingerprinting Scheme for Print-and-Capture Model / W-G. Kim [et al.] // Proc. 7th Pacific Rim Conf. on Multimedia, Advances in Multimedia Information Processing. – Hangzhou, China, 2006. – Vol. 4261. – P. 106–113.
10. Фергюсон, Н. Практическая криптография / Н. Фергюсон, Б. Шнайер. – М. : Изд. дом «Вильямс», 2005. – 416 с.
11. Зензин, О.С. Стандарт криптографической защиты – AES. Конечные поля / О.С. Зензин, М.А. Иванов ; под ред. М.А. Иванова – М. : КУДИЦ-Образ, 2002. – 176 с.
12. Панасенко, С.П. Алгоритмы шифрования. Специальный справочник / С.П. Панасенко. – СПб., 2009. – 576 с.
13. Singhal, N. Comparative Analysis of AES and RC4 Algorithms for Better Utilization / N. Singhal, J.P.S. Raina // Intern. Journal of Computer Trends and Technology. – 2011. – Vol. 1, iss. 1, no. 2. – P. 177–181.
14. Kocarev, L. Chaos-Based Cryptography. Theory, Algorithms and Applications / L. Kocarev, S. Lian. – Springer, 2011. – 395 p.
15. Otsu, N. A Threshold Selection Method from Gray-Level Histograms / N. Otsu // IEEE Transactions on Systems, Man, and Cybernetics. – 1979. – Vol. 9, no. 1. – P. 62–66.
16. Canny, J. A Computational approach to edge detection / J. Canny // IEEE Transactions on pattern analysis and machine intelligence. – 1986. – Vol. 8, no. 6. – P. 679–698.
17. Duda, R.O. Use of the Hough Transformation to Detect Lines and Curves in Pictures / R.O. Duda, P.E. Hart // Comm. ACM. – 1972. – Vol. 15. – P. 11–15.
18. Yang, H. Accurate localization of four extreme corners for barcode images captured by mobile phones / H. Yang, X. Jiang, A.C. Kot // Proc. of 17th Intern. Conf. on Image Processing (ICIP 2010). – Hong Kong, 2010. – P. 3897–3900.

19. Чехлов, Д.О. Нормализация изображений относительно перспективного преобразования на основе геометрических параметров / Д.О. Чехлов, С.В. Абламейко // Информатика. – 2004. – № 3. – С. 67–76.
20. Computer Vision Using Local Binary Patterns / M. Pietikäinen [et al.]. – London : Springer, 2011. – 207 p.
21. Толочко, А.М. Метод быстрого вычисления LBP-признаков области интереса видеоизображения / А.М. Толочко, А.А. Борискевич // Цифровая обработка сигналов. – 2012. – № 3. – С. 58–63.
22. Борискевич, А.А. Метод блочного маркирования изображений на основе синусоидальных решеток / А.А. Борискевич, Ю.А. Кочетков // Технические средства защиты информации : тез. докл. XI Белорусско-российской науч.-техн. конф., Минск, 2013 / Бел. гос. ун-т информатики и радиоэлектроники. – Минск, 2013. – С. 35.

Поступила 05.09.2016

*Белорусский государственный университет
информатики и радиоэлектроники,
Минск, П. Бровки, 6
e-mail: anbor@bsuir.by*

A.A. Boriskevich

EMBEDDING IDENTIFICATION DATA WITHIN IMAGES BASED ON SINUSOIDAL GRATINGS FOR MOBILE APPLICATIONS

A method of optically visualized block watermarking the images based on the proposed models of embedded message, marked image and robust textural correlation message extracting is developed. The results of computer modeling are presented.

УДК 004.942

В.М. Артемьев, А.О. Наумов, Л.Л. Кохан

ЛИНЕЙНАЯ ФИЛЬТРАЦИЯ МНОГОМЕРНЫХ СЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ РАСШИРЕННЫМ МЕТОДОМ НАИМЕНЬШИХ КВАДРАТОВ

При высокой размерности задачи фильтр Калмана становится труднореализуемым в реальном масштабе времени из-за больших вычислительных затрат. В качестве альтернативы рассматривается методика синтеза фильтров на основе метода наименьших квадратов по критерию минимума расширенной квадратичной невязки. Методика дает возможность сократить вычислительные затраты по нахождению коэффициентов усиления фильтра, однако при этом возрастает дисперсия ошибок фильтрации по сравнению с фильтром Калмана. На примере показывается степень этого увеличения за счет неучета априорной информации.

Введение

Для решения задач линейной фильтрации случайных последовательностей, наблюдаемых на фоне аддитивных помех, используют статистический подход на основе фильтра Калмана (ФК), обеспечивающий минимальную дисперсию ошибок фильтрации [1]. Однако при высокой размерности задачи ФК становится труднореализуемым в реальном масштабе времени из-за больших вычислительных затрат. Альтернативой статистическому служит детерминистский подход, когда фильтрация осуществляется на основе результатов лишь текущих измерений [2] с привлечением эмпирических данных. При этом снижение вычислительных затрат происходит за счет того, что не требуется проводить операции статистического усреднения. Для реализации такого подхода основным является метод наименьших квадратов (МНК), где в качестве критерия оптимальности используется квадратичная форма невязки решения [3, 4]. Получаемые при этом алгоритмы фильтрации требуют использования всей последовательности измерений с момента начала работы фильтра. Очевидно, что при большом числе периодов измерений фильтры такого рода становятся малоприменимыми. Задачу фильтрации целесообразно решать рекуррентно, как это имеет место у ФК, используя лишь оценки на предыдущем шаге фильтрации и текущие измерения. В такой постановке для синтеза алгоритмов фильтрации можно использовать расширенный метод наименьших квадратов (РМНК) [5], однако это обычно приводит к необходимости решения плохо обусловленных задач [6].

В настоящей работе предлагается методика синтеза рекуррентных линейных фильтров наименьших квадратов (ФНК) случайных последовательностей на основе РМНК, что позволяет решать широкий круг задач с меньшими вычислительными затратами, чем у ФК.

1. Формулировка задачи

Полезный сигнал \mathbf{x}_k является случайной последовательностью в виде m -мерного вектора $\mathbf{x}_k = [x_{1,k}, x_{2,k}, \dots, x_{m,k}]^T$, где $k = 0, 1, 2, \dots$ есть дискретное время, а символ « T » обозначает операцию транспонирования. Модель сигнала задается линейным стохастическим конечно-разностным уравнением вида

$$\mathbf{x}_k = A_k \mathbf{x}_{k-1} + \mathbf{w}_k, \quad (1)$$

где A_k – матрица размерности $m \times m$, а случайный m -мерный вектор $\mathbf{w}_k = [w_{1,k}, w_{2,k}, \dots, w_{m,k}]^T$ есть формирующее воздействие модели полезного сигнала в виде централизованного дискретного белого шума с ковариационной матрицей Q_k .

Модель n -мерного вектора измерений $\mathbf{z}_k = [z_{1,k}, z_{2,k}, \dots, z_{n,k}]^T$ определяется линейным уравнением

$$\mathbf{z}_k = H_k \mathbf{x}_k + \mathbf{v}_k, \quad (2)$$

где матрица измерений H_k имеет размерность $m \times n$, а случайный вектор $\mathbf{v}_k = [v_{1,k}, v_{2,k}, \dots, v_{n,k}]^T$ есть центрированный дискретный белый шум с ковариационной матрицей R_k , который моделирует ошибки измерений и является шумовым воздействием для этого канала.

Задача состоит в нахождении уравнений текущей оценки вектора полезного сигнала $\hat{\mathbf{x}}_k = [\hat{x}_{1,k}, \hat{x}_{2,k}, \dots, \hat{x}_{m,k}]^T$ на основе выбранного критерия оптимальности и текущих измерений \mathbf{z}_k методом РМНК. Это позволяет получить единую методику синтеза уравнений фильтрации для различных типов задач.

2. Уравнения фильтра

В основе методики синтеза лежит выбор критерия оптимальности $J_k(\hat{\mathbf{x}}_k)$. Для метода наименьших квадратов в качестве критерия используется квадратичная форма невязки (КФН) $(\mathbf{z}_k - H_k \hat{\mathbf{x}}_k)^T R_k^{-1} (\mathbf{z}_k - H_k \hat{\mathbf{x}}_k)$, где символ « $^{-1}$ » обозначает операцию обращения матрицы. Такой критерий зависит лишь от текущих оценок и не зависит от предыдущих, что не дает возможность получать рекуррентные оценки. В результате потребуется расширение КФН путем включения сглаживающей квадратичной формы, зависящей от предыдущих оценок. Выбор критерия оптимальности производится исходя из сущности решаемой задачи. Для сформулированных выше исходных данных подходящим вариантом является квадратичная форма $(\hat{\mathbf{x}}_k - A_k \hat{\mathbf{x}}_{k-1})^T Q_k^{-1} (\hat{\mathbf{x}}_k - A_k \hat{\mathbf{x}}_{k-1})$, которая следует из модели полезного сигнала (1). В результате критерий синтеза ФНК выглядит следующим образом:

$$J_k(\hat{\mathbf{x}}_k) = (1 - \alpha) (\mathbf{z}_k - H_k \hat{\mathbf{x}}_k)^T R_k^{-1} (\mathbf{z}_k - H_k \hat{\mathbf{x}}_k) + \alpha (\hat{\mathbf{x}}_k - A_k \hat{\mathbf{x}}_{k-1})^T Q_k^{-1} (\hat{\mathbf{x}}_k - A_k \hat{\mathbf{x}}_{k-1}), \quad (3)$$

где $0 < \alpha < 1$ является положительным весовым коэффициентом, величина которого выбирается из эмпирических соображений или по результатам моделирования и позволяет компенсировать отсутствие знаний априорной статистики. Этот критерий определяет текущие потери фильтрации в детерминистской постановке и его можно трактовать следующим образом: первое слагаемое учитывает влияние измерений \mathbf{z}_k на качество решения $\hat{\mathbf{x}}_k$; второе позволяет согласовать его с решением $\hat{\mathbf{x}}_{k-1}$ на предыдущем шаге фильтрации с учетом модели (1), а коэффициент α определяет вес этого слагаемого.

Оптимальные оценки $\hat{\mathbf{x}}_k$ находятся из условия минимума критерия (3). Необходимое условие минимума представляется в виде уравнения

$$\frac{\partial J_k(\hat{\mathbf{x}}_k)}{\partial \hat{\mathbf{x}}_k} = -2(1 - \alpha) H_k^T R_k^{-1} (\mathbf{z}_k - H_k \hat{\mathbf{x}}_k) + 2\alpha Q_k^{-1} (\hat{\mathbf{x}}_k - A_k \hat{\mathbf{x}}_{k-1}) = 0. \quad (4)$$

Его решение приводит к уравнению оптимального ФНК, которое имеет вид

$$\hat{\mathbf{x}}_k = K_{1,k} \hat{\mathbf{x}}_{k-1} + K_{0,k} \mathbf{z}_k. \quad (5)$$

Здесь матричный коэффициент $K_{1,k}$ размерности $m \times m$ имеет форму

$$K_{1,k} = \alpha (\alpha Q_k^{-1} + (1 - \alpha) H_k^T R_k^{-1} H_k)^{-1} Q_k^{-1} A_k \quad (6)$$

и задает экстраполяцию оценки $\hat{\mathbf{x}}_{k-1}$ на следующий шаг. Матричный коэффициент усиления $K_{2,k}$ размерности $m \times n$ определяется равенством

$$K_{0,k} = (1-\alpha) \left(\alpha Q_k^{-1} + (1-\alpha) H_k^T R_k^{-1} H_k \right)^{-1} H_k^T R_k^{-1} \quad (7)$$

и позволяет уточнить экстраполированное значение за счет наблюдения \mathbf{z}_k на текущем шаге фильтрации. Структура уравнения (5) говорит о рекуррентном характере фильтрации. Путем эквивалентных преобразований оно сводится к структуре фильтра с обратной связью:

$$\hat{\mathbf{x}}_k = K_{1,k}^* \hat{\mathbf{x}}_{k-1} + K_{0,k}^* \left(\mathbf{z}_k - H_k K_{1,k}^* \hat{\mathbf{x}}_{k-1} \right), \quad (8)$$

где

$$K_{1,k}^* = \left(I - K_{0,k} H_k \right)^{-1} K_{1,k} = A_k, \quad K_{0,k}^* = K_{0,k}. \quad (9)$$

В этом случае структура ФНК и первое слагаемое в формуле (8) совпадают с ФК.

Коэффициенты усиления ФНК находятся по формулам (6), (7) или (9), в то время как у ФК это требует решения ковариационного уравнения совместно с уравнением для оптимального коэффициента усиления, что более трудоемко. Однако уменьшение вычислительных затрат достигается за счет снижения точности фильтрации. Объем вычислений коэффициентов усиления ФНК в основном связан с обращением матрицы $(\alpha Q_k^{-1} + H_k^T R_k^{-1} H_k)$. Для вычислений можно воспользоваться формулой матричного тождества $(B^{-1} + C^T D^{-1} C)^{-1} = B - B C^T (D + C B C^T)^{-1} C B$ [7], что приводит к выражению

$$\left(\alpha Q_k^{-1} + (1-\alpha) H_k^T R_k^{-1} H_k \right)^{-1} = \alpha^{-1} L_k Q_k, \quad (10)$$

где

$$L_k = I - \alpha^{-1} (1-\alpha) Q_k H_k^T \left(R_k + \alpha^{-1} (1-\alpha) H_k Q_k H_k^T \right)^{-1} H_k.$$

3. Точность фильтрации

Качество работы ФНК оценивается величиной потерь (3), и для детерминистского подхода синтезированный фильтр оптимален. В то же время при статистическом подходе качество определяется путем оценки величин дисперсий ошибок фильтрации. В линейном случае оптимальным является ФК, обеспечивающий минимально возможные значения этих дисперсий. Очевидно, что в тех же условиях фильтрации дисперсии ошибок ФНК будут выше, поэтому результаты ФК могут служить оценкой нижних границ величин этих ошибок. Представляет интерес сравнение дисперсий ошибок обоих фильтров при одинаковых условиях работы, что дает возможность оценки снижения точности фильтрации при переходе от статистического подхода к детерминистскому.

Вектор ошибок фильтрации $\mathbf{e}_k = \mathbf{x}_k - \hat{\mathbf{x}}_k$. После подстановки выражений (1), (2), (5)–(7) приходим к следующему уравнению ошибок ФНК:

$$\mathbf{e}_k = A_k \mathbf{x}_{k-1} + \mathbf{w}_k - K_{1,k} \hat{\mathbf{x}}_{k-1} - K_{0,k} \mathbf{z}_k = K_{1,k} \mathbf{e}_{k-1} + L_k \mathbf{w}_k - K_{0,k} \mathbf{v}_k. \quad (11)$$

Вектор математических ожиданий ошибок обозначается через $\mathbf{m}_{e,k} = \langle \mathbf{e}_k \rangle$, где угловые скобки – это операция нахождения математического ожидания. Усредняя обе части (11), получаем уравнение для вектора математических ожиданий ошибок:

$$\mathbf{m}_{e,k} = K_{1,k} \mathbf{m}_{e,k-1} + L_k \mathbf{m}_{w,k} - K_{0,k} \mathbf{m}_{v,k}, \quad (12)$$

где $\mathbf{m}_{w,k} = \langle \mathbf{w}_k \rangle$ и $\mathbf{m}_{v,k} = \langle \mathbf{v}_k \rangle$ есть математические ожидания воздействий. В рассматриваемом случае полагаем их равными нулю, что сводит (12) к однородному уравнению $\mathbf{m}_{e,k} = K_{1,k} \mathbf{m}_{e,k-1}$. Его решение с течением времени стремится к нулю, и это говорит о том, что уравнение ФНК (5) обеспечивает несмещенность оценок.

При статистически независимых дискретных белых шумах \mathbf{w}_k и \mathbf{v}_k и с учетом независимости значений ошибок \mathbf{e}_{k-1} в момент времени $k-1$ от воздействий \mathbf{w}_k и \mathbf{v}_k в последующий момент времени k можно получить уравнение для ковариационной матрицы ошибок $P_k = \langle (\mathbf{e}_k - \mathbf{m}_{e,k})(\mathbf{e}_k - \mathbf{m}_{e,k})^T \rangle$ в виде

$$P_k = K_{1,k} P_{k-1} K_{1,k}^T + L_k Q_k L_k^T + K_{0,k} R_k K_{0,k}^T \quad (13)$$

с начальным условием P_0 . Диагональные элементы этой матрицы являются дисперсиями ошибок фильтрации каждой из составляющих вектора \mathbf{x}_k .

Выбор величины коэффициента α можно осуществлять следующим образом. Первоначально задается средним значением $\alpha = 0,5$, что соответствует критерию максимального правдоподобия, если воздействия \mathbf{w}_k и \mathbf{v}_k имеют гауссово распределение [4]. Полученные из уравнения (13) значения дисперсий ошибок будут больше, чем у ФК. Затем путем моделирования находят значения α , соответствующие устойчивому решению (13) и наименьшим значениям дисперсий ошибок. В итоге ФНК обеспечивает величину дисперсий, хотя и большую, чем у ФК, но меньшую, чем у метода максимального правдоподобия.

4. Пример сравнения дисперсий ошибок ФНК и ФК

Рассмотрим задачу оптимальной фильтрации текущих координат дальности до маневрирующего воздушного объекта по данным радиолокатора сопровождения. Статистическая модель изменения координат приведена в работе [8] и имеет вид системы стохастических конечно-разностных уравнений:

$$\begin{aligned} x_{1,k} &= x_{1,k-1} + T_0 x_{2,k-1} + 0,5T_0^2 x_{3,k-1}; \\ x_{2,k} &= x_{2,k-1} + T_0 x_{3,k-1}; \\ x_{3,k} &= a_3 x_{3,k-1} + w_{3,k}; \quad 0 < a_3 < 1. \end{aligned} \quad (14)$$

В уравнениях (14) посредством $x_{1,k}$ обозначена дальность до объекта (м) в моменты времени $k = 0, 1, 2, \dots$; $x_{2,k}$ – скорость изменения дальности (м/с); $x_{3,k}$ – ускорение (м/с²), которое формируется воздействием $w_{3,k}$ в виде центрированного дискретного белого шума с постоянной дисперсией $\sigma_{w_3}^2$. Параметр T_0 (с) задает величину периода поступления измерений дальности от радиолокатора сопровождения. В векторной форме уравнения (14) могут быть представлены в виде (1) со следующими значениями векторов и матриц:

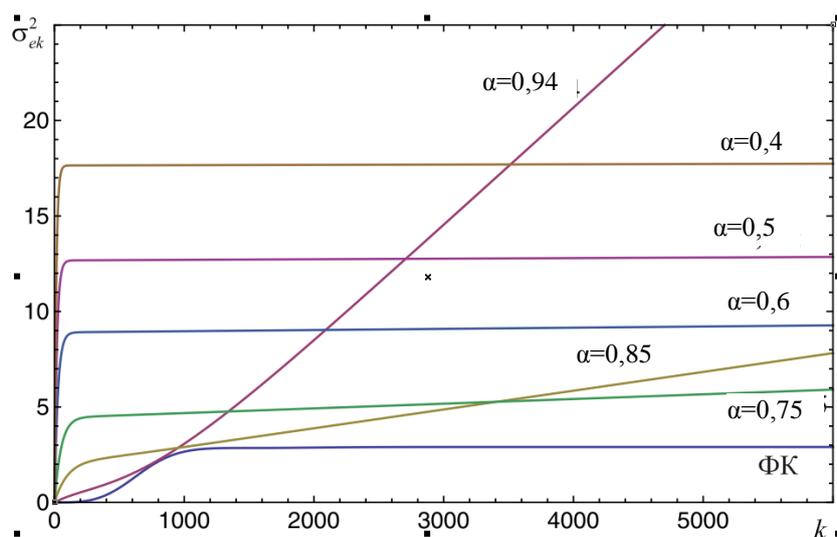
$$\mathbf{x}_k = \begin{pmatrix} x_{1,k} \\ x_{2,k} \\ x_{3,k} \end{pmatrix}; \quad A = \begin{pmatrix} 1 & T_0 & 0,5T_0^2 \\ 0 & 1 & T_0 \\ 0 & 0 & a_3 \end{pmatrix}; \quad \mathbf{w}_k = \begin{pmatrix} 0 \\ 0 \\ w_{3,k} \end{pmatrix}; \quad Q_k = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sigma_{w_3}^2 \end{pmatrix}. \quad (15)$$

Полагаем, что дальность $x_{1,k}$ измеряется с аддитивной ошибкой $v_{1,k}$ в виде центрированного белого шума с постоянной дисперсией $\sigma_{v_1}^2$, статистически независимого от $w_{3,k}$. Линейная модель измерений задается уравнением (2), где имеют место оба значения:

$$H = (1 \quad 0 \quad 0); \quad \mathbf{z}_k = (z_{1,k}); \quad \mathbf{v}_k = (v_{1,k}); \quad R_k = (\sigma_{v_1}^2). \quad (16)$$

При расчетах использованы следующие исходные данные: период поступления измерений $T_0 = 10^{-3}$ с; дисперсия ошибки измерения дальности $\sigma_{v_1}^2 = 900$ м²; дисперсия изменения ускорения в установившемся режиме $\sigma_3^2 = 400$ м²/с⁴; величина коэффициента $a_3 = 0,968$, соответствующая длительности корреляции ускорения в установившемся режиме 30 с; дисперсия формирующего воздействия ускорения для этих данных $\sigma_{w_3}^2 = 25,2$ м²/с⁴.

С помощью выражения (13) для ковариационной матрицы ошибок ФНК и ковариационного уравнения ФК [1, 2] при заданных параметрах модели входного воздействия и ряда значений коэффициентов α получены графики изменения дисперсий ошибок оценок дальности σ_{ek}^2 у ФК и ФНК (рисунок).



Дисперсии ошибок измерения дальности σ_{ek}^2 у ФК и ФНК при различных значениях весового коэффициента α

Графики показывают, что наименьшими значениями дисперсий ошибок обладает ФК. Для метода максимального правдоподобия ($\alpha = 0,5$) эта дисперсия возрастает примерно в четыре раза. При увеличении значений весового коэффициента α дисперсия ошибок вначале уменьшается (до $\alpha = 0,8$), а затем решение становится неустойчивым ($\alpha = 0,85; 0,94$). Наименьшее значение устойчивой величины дисперсии имеет место при $\alpha = 0,75$, когда дисперсия ошибок ФНК примерно в два раза больше, чем у ФК.

Заключение

Использование ФНК на основе расширенного метода наименьших квадратов позволяет сократить затраты на вычисление коэффициентов усиления по сравнению с ФК, однако это достигается за счет увеличения дисперсий ошибок фильтрации. Снижение этих величин осуществляется путем эмпирического выбора весового коэффициента, что делает возможным получение дисперсий ошибок ФНК, хотя и больших, чем у ФК, но меньших, чем при использовании метода максимального правдоподобия.

Список литературы

1. Jazwinski, A.H. Stochastic Process and Filtering Theory / A.H. Jazwinski. – N.Y. : Academic Press, 1970.– 362 p.
2. Степанов, А.О. Основы теории оценивания с приложениями к задачам обработки навигационной информации : в 3 ч. / А.О. Степанов. – СПб. : ГНЦ РФ ЦНИИ «Электроприбор», 2009. – Ч. 1. : Введение в теорию оценивания. – 496 с.

3. Линник, Ю.В. Метод наименьших квадратов и основы математико-статистической теории обработки наблюдений / Ю.В. Линник. – Изд. 2-е, доп. и испр. – М. : Физматгиз, 1962. – 349 с.
4. Эльясберг, П.Е. Определение движения по результатам измерений / П.Е. Эльясберг. – М. : Наука, 1976. – 267 с.
5. Альберт, А. Регрессия, псевдоинверсия и рекуррентное оценивание / А. Альберт. – М. : Наука, 1977. – 215 с.
6. Тихонов, А.Н. Методы решения некорректных задач / А.Н. Тихонов, В.Я. Арсенин. – М. : Наука, 1979. – 232 с.
7. Себер, Д. Линейный регрессионный анализ / Д. Себер. – М. : Мир, 1980. – 322 с.
8. Фарина, А. Цифровая обработка радиолокационной информации. Сопровождение целей / А. Фарина, Ф. Студер. – М. : Радио и связь, 1993. – 319 с.

Поступила 07.10.2016

*Институт прикладной физики
НАН Беларуси,
Минск, Академическая, 16
e-mail: naumov@iapf.bas-net.by*

V.M. Artemiev, A.O. Naumov, L.L. Kokhan

LINEAR FILTRATION OF RANDOM SEQUENCES USING A LEAST SQUARE METHOD WITH REGULARIZATION

At high problem dimension the Kalman filter becomes difficult to realize in real time due to the high computational costs. Alternatively, a technique of filter synthesis on the basis of the extended least square method with extended square discrepancy is given. The technique allows to reduce the computation costs for the search of filter gain coefficients, but it increases the variance of the filtering error in comparison with Kalman filter. The degree of this increase in case of prior information is not taken into account is shown on example.

К 80-ЛЕТИЮ ЧЛЕНА-КОРРЕСПОНДЕНТА Г.В. РИМСКОГО

УДК 683.735.33

В.А. Вишняков**НАУЧНОЕ НАСЛЕДИЕ ЧЛЕНА-КОРРЕСПОНДЕНТА Г.В. РИМСКОГО**

Представлена научная деятельность профессора Г.В. Римского с 1960 по 2000 г., разделенная на четыре периода. Приведены основные научные результаты этих периодов.

Введение

Автору довелось быть одним из первых учеников Геннадия Васильевича Римского, на протяжении 30 лет работать под его руководством, а в дальнейшем создать свою научную школу по информационному управлению в Минском инновационном университете (МИУ) и Белорусской государственной академии связи.

В начале 1970-х гг. на кафедре ЭВМ Минского радиотехнического института под руководством доцента Г.В. Римского образовалась научная группа молодых аспирантов, названная GRLT (General Roots Locus Theory). В нее вошли В.А. Вишняков, Ю.А. Скудняков, В.П. Соловьев, В.В. Таборовец. В рамках этой группы разрабатывались как теоретические, так и практические основы общей теории корневых траекторий. В результате были разработаны и внедрены комплексы программ для автоматизации управления траловым ловом рыбы (Специальное экспериментально-конструкторское бюро промышленного рыболовства), автоматизации управления системами прокатных станков (Челябинский металлургический завод и Уральский НИИ трубной промышленности). Комплекс был включен в состав САПР САУ Минвуза РСФСР и в учебный процесс ряда вузов страны. В период с 1978 по 1980 г. членами группы защищены четыре кандидатские диссертации, а руководителем – докторская диссертация в МВТУ им. Баумана.

Научную деятельность профессора Г.В. Римского можно разбить на четыре периода:

1959–67 гг. – разработка основ общей теории корневых траекторий и некоторых ее приложений [1];

1968–80 гг. – разработка общей теории корневых траекторий и средств автоматизации проектирования систем автоматического управления на ее основе [2, 3];

1981–90 гг. – развитие общей теории корневых траекторий и разработка приложений для исследования и проектирования технологических систем и систем автоматизации проектирования [4, 5];

1991–2000 гг. – развитие общей теории корневых траекторий и разработка приложений для исследования интервальных динамических систем [6–8].

Основные научные результаты первого периода (1959–1967)

Сформулированы основные принципы общей теории корневых траекторий: отображение произвольных кривых, их совокупностей и заданных областей комплексных плоскостей свободных параметров динамических систем на плоскости комплексных переменных; принцип последовательных многократных преобразований плоскостей свободных параметров и принцип функциональной параметризации.

Разработаны методы построения, анализа и синтеза корневых годографов произвольных образов, позволяющие находить пределы изменения характеристик качества САУ, их критические режимы работы, синтезировать корневые годографы компьютерными средствами, что дает возможность аналитического и графического исследования корневой чувствительности САУ.

Разработаны методы определения корневых годографов различных типов для многосвязных систем, систем с запаздыванием, а также способы аналитического описания корневых годографов в логарифмических системах координат.

На основе общей теории корневых траекторий предложены и разработаны корневые методы решения задач абсолютной устойчивости положения равновесия нелинейных систем управления.

Основные научные результаты второго периода (1968–1980)

Разработаны основы общей теории корневых траекторий (ОТКТ) САУ и сформулированы основные ее принципы: принцип отображения произвольных кривых, их совокупностей и заданных областей комплексных плоскостей свободных параметров на плоскость комплексных переменных и принцип последовательных многократных преобразований свободных плоскостей.

Разработаны методы математического описания корневого годографа, полей корневых траекторий, способы преобразования уравнений траекторий корней различных типов, функций корневой чувствительности. Рассмотрены методы построения, анализа и синтеза корневых годографов произвольного образа. Предложены и развиты методы построения заданных областей в пространстве коэффициентов функций отображения. Математический аппарат многомерных коммутационных преобразований числовых массивов изоморфных функций от нескольких независимых переменных обеспечивает однородность вычислительных процедур, что упрощает их алгоритмизацию и реализацию на компьютерах, и является основой для разработки машинно ориентированных корневых методов расчета и исследования САУ.

Предложен и разработан новый класс точных методов исследования устойчивости нелинейных систем на основе ОТКТ. В отличие от частотных методов и методов, основанных на использовании функции Ляпунова, данные методы базируются на анализе поведения ветвей корневых годографов специальных типов в комплексных плоскостях относительно координатных осей. С помощью корневых методов решены основные задачи абсолютной устойчивости положения равновесия нелинейных САУ, асимптотической устойчивости в большом, гиперустойчивости, L_2 -устойчивости, K -устойчивости и абсолютной устойчивости процессов.

Получены критерии абсолютной устойчивости в пространстве параметров нелинейных САУ и разработаны способы построения границ областей устойчивости и заданного качества. Корневые методы не дают точечные решения, а позволяют находить допустимые области в пространстве критериев и параметров САУ. Отличительными особенностями корневых методов являются однородность вычислительных процедур и возможность аналитических преобразований уравнений динамики с помощью компьютеров, что позволяет свести к минимуму влияние ограничений разрядности, итерационных расчетов, использовать компьютеры не только для проектирования САУ, но и для теоретических исследований решения задач в общем виде. Машинная ориентация корневых методов расчета САУ обеспечивает простоту их программной реализации, расширяет область практического применения ОТКТ.

Разработаны методы и средства автоматизации исследования систем уравнений на основе ОТКТ. Отличительной особенностью предложенных методов является широкое использование аналитических преобразований функций нескольких независимых переменных на компьютерах, что позволяет реализовать вывод уравнений траекторий корней, границ областей заданного качества САУ в символьном виде на компьютерах и построить комплекс программ GRLT для расчета и исследования САУ на языках программирования.

Разработаны структуры вычислительных комплексов для автоматизированного проектирования (КАПР) систем управления, рассмотрены принципы построения однородных комплексов, способы организации вычислительных процессов автоматизированного проектирования. Существенным отличием предложенной организации КАПР является интеграция технических и программных средств, обеспечивающих коллективную взаимозависимую работу разработчиков над одним проектом, оперативное управление процессом проектирования.

Основные научные результаты третьего и четвертого периодов (1981–2000)

Разработан метод многонастраиваемых параметров многорежимных линейных интервальных динамических систем (ЛИДС), отличающихся использованием ортогональных корневых годографов для построения предельно допустимого положения доминирующих корней ха-

рактических уравнений каналов, при которых обеспечивается получение требуемых динамических свойств. Доказаны необходимые и достаточные условия относительной устойчивости многомерных ЛИДС. Разработан метод обеспечения апериодичности ЛИДС.

Предложен метод определения запаса устойчивости ЛИДС с аддитивной и мультипликативной неопределенностью. Предложен метод управления миграцией корней характеристических уравнений ЛИДС с дисковыми коэффициентами. Найдены необходимые и достаточные условия размещения корней характеристического уравнения данных ЛИДС в областях заданного качества.

Разработаны метод построения математических моделей адаптивных технологий автоматизированного проектирования интервальных динамических систем на базе семантического полигона и способ идентификации динамических характеристик процессов.

Предложен способ упреждающего управления технологиями проектирования ЛИДС на основе прогнозирования процессов на семантическом полигоне и на основе эталонных моделей.

Разработаны алгоритмы построения передаточных функций каналов многомерных ЛИДС и структурных преобразований их блок-схем в символьной форме, в основе которых лежит описание структурных схем в табличной форме. Предложены алгоритмы построения и модификации семантического полигона, объекты и способы обмена информацией.

Предложены структура и организация функционирования визуальной системы автоматизированного проектирования ЛИДС. Разработаны основные подсистемы визуальной инструментальной системы поддержки проектирования, отличающиеся визуализацией процедур построения структурных схем ЛИДС, вычислительных моделей интервальных систем средствами адаптивного управления проектированием.

Разработаны основы построения интеллектуальных комплексов автоматизированного проектирования, отличающихся коллективной работой пользователей, гибкими технологиями проектирования, интеллектуальными решателями задач.

Заключение

После защиты докторской диссертации Г.В. Римский получил звание профессора, возглавлял кафедру КиП ЭВА в Минском радиотехническом институте, а в 1986 г. перешел в Институт технической кибернетики АН БССР на должность заведующего лабораторией. С 1992 г. работал в ВАК Республики Беларусь, где возглавлял отдел технических и военных наук. В 1995 г. Г.В. Римский стал лауреатом премии НАН Беларуси за цикл монографий по общей теории корневых траекторий систем управления и автоматизации их проектирования. В 1996 г. избран членом-корреспондентом НАН Беларуси. Г.В. Римский был признан в сфере науки «Человеком года 1997» в республике. Им создана научная школа из 19 кандидатов и 2 докторов наук, он является автором более 330 научных работ, в том числе 7 монографий, 30 изобретений, 13 патентов на промышленные образцы. Во время работы в ВАК он организовал сеть экспертных советов и их взаимодействие с коллегами из России, Украины и из других стран СНГ, подготовил ряд важнейших аналитических и нормативных документов по аттестации научно-педагогических кадров высшей квалификации в Республике Беларусь.

Автор статьи в 1981–82 гг. прошел научную стажировку в Датском техническом университете по теме «Системы автоматического проектирования и управления», создал свою научную группу по разработке специализированных ЭВМ, в 1990 г. в Институте кибернетики им. Глушкова защитил докторскую диссертацию в области объектно-ориентированных и логических систем. В 1992 г. получил аттестат профессора. В 1996 г. прошел стажировку в США в области телекоммуникаций и информационных технологий в бизнесе и получил три сертификата. В 1997 и 2001 гг. выполнял научные работы в Германии по линии DAAD в области интернет-технологий в бизнесе. В 1996–2002 гг. создал и возглавил кафедру экономики и управления ВГКС, в 2002–2011 гг. – кафедру менеджмента в МИУ. Им опубликовано более 340 научных работ, в том числе 5 монографий, 30 учебных пособий (4 с грифом Министерства образования Республики Беларусь), более 100 статей; подготовлены 5 кандидатов технических наук и 20 магистров. В настоящее время работает профессором МИУ и по совместительству профессором БГУИР.

Список литературы

1. Римский, Г.В. Основы общей теории корневых траекторий систем автоматического управления / Г.В. Римский. – Минск : Наука и техника, 1972. – 328 с.
2. Римский, Г.В. Автоматизация анализа динамических систем на основе общей теории корневых траекторий / Г.В. Римский. – Минск : Ин-т техн. кибернетики АН БССР, 1973. – 261 с.
3. Римский, Г.В. Автоматизация исследований динамических систем / Г.В. Римский, В.В. Таборовец. – Минск : Наука и техника, 1978. – 336 с.
4. Аналитическое преобразование на ЭВМ структур систем управления : материалы по математическому обеспечению ЭВМ : в 2 ч. / авт.-сост. : Г.В. Римский, В.В. Таборовец, И.П. Стацук. – Минск : Ин-т техн. кибернетики АН БССР, 1988. – Ч. 1. – 158 с., Ч. 2. – 146 с.
5. Римский, Г.В. Теория систем автоматизированного проектирования: интеллектуальные САПР на базе вычислительных комплексов и сетей / Г.В. Римский. – Минск : Наука и техника, 1994. – 631 с.
6. Римский, Г.В. Распределение комплекса взаимосвязанных работ на многопроцессорной вычислительной системе / Г.В. Римский, С.В. Чебаков. – Минск : Ин-т техн. кибернетики НАН Беларуси, 1998. – 16 с.
7. Римский, Г.В. Фразеологическая информационно-поисковая система литературного применения / Г.В. Римский, О.О. Бабинец. – Минск, 1998. – 64 с. – (Препринт / Нац. акад. наук Беларуси, Ин-т техн. кибернетики; № 4).
8. Корневые методы исследования интервальных систем / Г.В. Римский [и др.]. – Минск : Ин-т техн. кибернетики НАН Беларуси, 1999. – 186 с.
9. Поэт в науке, учёный в поэзии (к 75-летию со дня рождения Г.В. Римского) / А.В. Тузиков [и др.] // Информатика. – 2011. – № 3. – С. 129–132.

Поступила 05.10.2016

*Минский инновационный университет,
Минск, ул. Лазо, 12
e-mail: vish2002@list.ru*

U.A. Vishniakou

SCIENTIFIC LEGACY OF PROFESSOR G.V. RIMSKIY

Four periods of research activity of professor G.V. Rimskiy in 1960–2000 are presented. The main research results of these periods are shown.

УДК 681.51.033.26

А.А. Несенчук

ИССЛЕДОВАНИЕ ПОВЕДЕНИЯ И СИНТЕЗ СЕМЕЙСТВ ХАРАКТЕРИСТИЧЕСКИХ ПОЛИНОМОВ ИНТЕРВАЛЬНЫХ ДИНАМИЧЕСКИХ СИСТЕМ НА ОСНОВЕ ИСПОЛЬЗОВАНИЯ ФУНКЦИИ ПАРАМЕТРА КОРНЕВОГО ПОРТРЕТА СИСТЕМЫ

Анализируется поведение динамических систем четвертого порядка в условиях интервальной вариации параметров с использованием корневых портретов. Выполняется исследование динамики и устанавливаются закономерности поведения корневых портретов систем четвертого порядка с интервальной неопределенностью на границе асимптотической устойчивости. Определяются условия устойчивости и выражения для вычисления интервалов параметров, при которых обеспечивается робастная устойчивость системы.

Введение

Задачи анализа и синтеза управления в динамических системах различной физической природы (в том числе технических), функционирующих в условиях значительной параметрической неопределенности, в настоящее время весьма актуальны [1–8]. Одной из них является, например, задача управления потокосцеплением в системах векторного управления электроприводом в условиях параметрической неопределенности, поскольку качество управления потокосцеплением в значительной степени определяет качество управления электромагнитным моментом и скоростью, а также энергетическую эффективность привода. Поэтому большое значение имеет исследование устойчивости и параметрического синтеза динамических систем управления объектами, параметры которых изменяются в некоторых интервалах значений.

Среди современных методов синтеза систем управления сложными техническими объектами [5] наиболее распространенными являются корневые [1, 2, 6], частотные [7, 8], алгебраические [3, 4], модального управления [5], позволяющие задавать желаемое расположение корней в плоскости собственных частот, пространства состояний [5], хорошо подходящие для структурного синтеза, и ряд других. Интересны задачи об устойчивости в условиях неопределенности, решаемые в современных постановках в робастном варианте [2–8], для которых может быть применен корневой подход. Преимущество данного подхода к проблеме состоит в том, что само его использование уже предполагает параметрические вариации (неопределенность). Он является мощным методом синтеза систем и отличается большой наглядностью, позволяя не только рассчитывать требуемые значения параметров системы, но и в деталях наблюдать характер изменения динамических свойств и реакцию системы в ответ на параметрические вариации, что особенно важно при исследовании систем с неопределенными, в частности интервальными, параметрами.

В настоящей статье рассматривается метод исследования и параметрического синтеза динамических систем четвертого порядка с интервальной неопределенностью, основанный на установлении закономерностей поведения корневых портретов систем на границе асимптотической устойчивости.

1. Корневой портрет интервальной динамической системы

Рассмотрим динамическую систему, описываемую семейством интервальных характеристических уравнений [3] четвертого порядка вида

$$s^4 + a_1s^3 + a_2s^2 + a_3s + a_4 = 0, \quad (1)$$

где s – комплексное переменное, $s = \sigma + i\omega$.

Изменение коэффициентов характеристического уравнения динамической системы в определенных интервалах отражает вариацию ее физических параметров, их нежелательное изменение по отношению к номинальным значениям, а именно параметрическую неопределенность системы. Предположим, что коэффициенты (1) действительны, неотрицательны и изменяются в следующих пределах:

$$\underline{a}_j \leq a_j \leq \bar{a}_j, \quad j = \overline{1,4}, \quad a_0 = 1, \quad (2)$$

где \underline{a}_j и \bar{a}_j – соответственно минимальное и максимальное значения замкнутого интервала изменения значений коэффициента a_j .

Устойчивость системы с параметрической неопределенностью [5–7], описываемой семейством характеристических уравнений (1), определяется расположением корней семейства относительно мнимой оси $i\omega$ плоскости корней s , являющейся границей асимптотической устойчивости системы.

Перепишем уравнение (1), сделав в нем замену переменного $s = \sigma + i\omega = i\omega$ ($\sigma = 0$):

$$\omega^4 - a_1\omega^3i - a_2\omega^2 + a_3\omega i + a_4 = 0, \quad (3)$$

и на основе (3) выведем уравнение корневого годографа на границе устойчивости

$$-a_1\omega^3 + a_3\omega = 0 \quad (4)$$

и уравнение (функцию) параметра [4] на границе устойчивости

$$f(\omega) = -\omega^4 + a_2\omega^2 = a_4 \quad (5)$$

для рассматриваемой интервальной динамической системы (ИДС) [5–7].

Под корневым годографом будем понимать *свободный корневой годограф* [6], т. е. формируемый при вариации свободного члена уравнения (1).

Уравнение корневого годографа (4) представляет собой также уравнение линий уровня поля F_i свободных корневых траекторий [6], параметром которого [6] определим коэффициент $a_{n-1} = a_3$. Оно используется для нахождения координат ω точек пересечения границы устойчивости (оси $i\omega$) ветвями корневых годографов данного поля и *корневого портрета* (годографа)

$$P = \{F_i \mid i = 1, 2, \dots\} \quad (6)$$

системы [6] в целом, который представляет собой бесконечное множество полей корневых траекторий и, следовательно, является носителем их свойств.

Функция параметра (5) применяется для определения значений параметра a_4 , который называется *свободным параметром* годографа [6].

2. Область пересечений корневого портрета системы

Ввиду аналитичности и непрерывности функций (4) и (5) очевидно, что точки пересечения границы устойчивости (оси $i\omega$) ветвями годографов семейства P полей корневых траекторий (6) при изменении коэффициентов (1) в пределах интервала значений

$$0 < a_j < +\infty \quad (7)$$

образуют на этой оси некоторую *область пересечений* D_ω^P .

Определение 1. Областью пересечений D_ω^P корневого портрета интервальной динамической системы, описываемой характеристическим уравнением (1), назовем область $[0, +\infty]$ на

границе $i\omega$ асимптотической устойчивости системы, в пределах которой располагаются значения функции параметра (5) для данного портрета при условии вариации всех коэффициентов a_j уравнения (1) в пределах (7).

Очевидно, что D_ω^P представляет собой непрерывную область. Полная область пересечений охватывает также и отрицательную полуось оси $i\omega$, но ввиду симметричности портрета относительно оси σ здесь рассматривается только верхняя полуплоскость плоскости корней s .

Каждое поле F_i (6) и каждая ветвь b_i годографов поля образуют некоторые подобласти: подобласть $D_{\omega_i}^F$, $i = 1, 2, \dots$, и непрерывную подобласть $D_{\omega_i}^b$, $i = 1, 2, \dots$, указанной области D_ω^P соответственно.

3. Область экстремумов корневого портрета системы. Мажоранта и миноранта

Определим максимальные и минимальные значения функции параметра (5) в подобласти $D_{\omega_i}^F \subset D_\omega^P$, образуемой годографами поля корневых траекторий (6) при пересечении границы устойчивости. Для этого выполним исследование данной функции на экстремум с помощью первой производной.

Очевидно, что максимальное значение параметра a_4 достигается при $a_2 = \bar{a}_2$, а минимальное значение – при $a_2 = \underline{a}_2$. Перепишем (5) для определения мажоранты функции параметра:

$$a_{4\max} = -\omega^4 + \bar{a}_2\omega^2. \quad (8)$$

Возьмем первую производную от функции параметра (8) и приравняем ее нулю:

$$-4\omega^3 + 2\bar{a}_2\omega = 0. \quad (9)$$

Решив (9), получим три точки экстремума для мажоранты функции параметра на границе устойчивости для поля при $a_2 = \bar{a}_2$:

$$\begin{aligned} \omega_{e\max} &= 0, & a_{4e\max} &= 0; \\ \omega_{e\max} &= \pm \sqrt{\frac{\bar{a}_2}{2}}, & a_{4e\max} &= -\omega_{e\max}^4 + \bar{a}_2 \cdot \omega_{e\max}^2. \end{aligned} \quad (10)$$

Очевидно, что мажоранта (8) с точками экстремума (10) является также мажорантой всего корневого портрета в целом.

Перепишем (5) для определения миноранты функции параметра:

$$a_{4\min} = -\omega^4 + \underline{a}_2\omega^2. \quad (11)$$

Возьмем первую производную от функции параметра (11) и приравняем ее нулю:

$$-4\omega^3 + 2\underline{a}_2\omega = 0. \quad (12)$$

Решив (12), получим три точки экстремума для миноранты функции параметра на границе устойчивости для поля при $a_2 = \underline{a}_2$:

$$\begin{aligned} \omega_{e\min} &= 0, & a_{4e\min} &= 0; \\ \omega_{e\min} &= \pm \sqrt{\frac{\underline{a}_2}{2}}, & a_{4e\min} &= -\omega_{e\min}^4 + \underline{a}_2 \cdot \omega_{e\min}^2. \end{aligned} \quad (13)$$

Для упрощения рисунка точки экстремума функции соединены прямыми линиями, хотя в действительности они представляют собой кривые.

Значения $\omega_{e_{\min}}$ и $\omega_{e_{\max}}$ представляют собой координаты ω точек, соответствующих минимальному и максимальному значениям экстремума функции параметра корневого портрета интервальной динамической системы.

Таким образом, область пересечений D_{ω}^P может быть разделена на три области (см. рис. 1):

D_{ω}^+ – возрастания функции параметра;

D_{ω}^- – убывания функции параметра;

D_{ω}^c – сочетания участков возрастания и убывания функции параметра.

Интервал $[z', z''] \subseteq Z_{\omega}$ на рис. 1 представляет собой область, в которой происходит миграция начальных точек (полюсов) [б] корневого портрета через границу устойчивости в правую полуплоскость корней. Точкам z', z'' соответствуют точки z_1, z_2 на диаграмме.

На интервале $[0, z']$, охватывающем полностью область D_{ω}^+ и частично область D_{ω}^c :

$$D_{\omega}^+ \subset [0, z'], [0, z'] \cap D_{\omega}^c,$$

граница устойчивости пересекается только положительными ветвями семейства годографов корневого портрета системы, так как значения функции параметра (5) здесь только положительны (см. рис. 1). Семейство начальных точек годографов корневого портрета системы полностью располагается в левой полуплоскости корней, однако положительные ветви годографов портрета частично лежат в правой полуплоскости корней. Следовательно, в семействе (1) полиномов, описывающих динамику системы, могут быть неустойчивые полиномы, но при определенных значениях интервалов (2) семейство полностью устойчиво. Это область устойчивости системы.

На интервале $[z', z'']$, частично принадлежащем области D_{ω}^c и частично области D_{ω}^- :

$$[z', z''] \cap D_{\omega}^c, [z', z''] \cap D_{\omega}^-,$$

граница устойчивости пересекается как положительными, так и отрицательными ветвями семейства годографов корневого портрета системы, так как функция параметра (5) принимает здесь как положительные, так и отрицательные значения (см. рис. 1). Некоторые начальные точки, а следовательно, и целые ветви годографов корневого портрета мигрировали в правую полуплоскость, т. е. в семействе (1) в данном случае содержатся как устойчивые, так и неустойчивые полиномы. Эта область представляет собой область неустойчивости системы.

На интервале $[z'', \infty]$, полностью принадлежащем области D_{ω}^- :

$$[z'', \infty] \subset D_{\omega}^-,$$

граница устойчивости пересекается только отрицательными ветвями семейства годографов корневого портрета системы, так как значения функции параметра (5) здесь только отрицательны (см. рис. 1). Все семейство начальных точек годографов, а следовательно, и соответствующих положительных ветвей корневого портрета в данном случае мигрировало в правую полуплоскость. При этом среди ветвей годографов корневого портрета системы не имеется ни одной положительной ветви (части ветви), расположенной в левой полуплоскости, а в семействе (1), следовательно, не имеется ни одного устойчивого полинома. Эта область представляет собой область неустойчивости, и данное состояние портрета характеризует состояние неустойчивости системы.

5. Реальная область пересечений

Определим пределы области D_{ω}^R пересечения границы устойчивости ветвями корневого портрета интервальной динамической системы, описываемой характеристическим полиномом (1). С этой целью рассмотрим уравнение корневого годографа системы (4) и определим значения его корней:

$$-a_1\omega^3 + a_3\omega = 0;$$

$$\begin{cases} \omega = 0, \\ \omega^2 = \frac{a_3}{a_1}; \end{cases} \quad \begin{cases} \omega = 0, \\ \omega = \pm \sqrt{\frac{a_3}{a_1}}. \end{cases}$$

Очевидно, что при $\omega > 0$

$$\omega_{\max} = \sqrt{\frac{\bar{a}_3}{\underline{a}_1}}, \quad \omega_{\min} = \sqrt{\frac{\underline{a}_3}{\bar{a}_1}}, \quad (14)$$

где ω_{\max} и ω_{\min} – границы реальной области пересечений D_{ω}^R границы устойчивости ветвями годографов корневого портрета в целом.

Определение 3. Реальной областью пересечений D_{ω}^R корневого портрета интервальной динамической системы, описываемой характеристическим уравнением (1), назовем область $[\omega_{\min}, \omega_{\max}]$ (14) на границе $i\omega$ асимптотической устойчивости системы, в пределах которой имеет место пересечение этой границы ветвями годографов корневого портрета:

$$[\omega_{\min}, \omega_{\max}] \subseteq D_{\omega}^R. \quad (15)$$

Очевидно, что границы (14) реальной области пересечений D_{ω}^R определяются значениями коэффициентов a_1 и a_3 в отличие от границ области экстремумов D_{ω}^e , которые определяются значениями коэффициента a_2 (см. (10), (13)).

6. Условия устойчивости системы и определение параметров, обеспечивающих устойчивость

Сформулируем условия устойчивости интервальной динамической системы четвертого порядка, описываемой семейством характеристических полиномов (1). Эти условия зависят от того, каким образом располагается реальная область пересечений D_{ω}^R корневого портрета системы по отношению к областям D_{ω}^+ возрастания функции параметра, D_{ω}^- убывания функции параметра и D_{ω}^c сочетания участков возрастания и убывания функции параметра. Определим три возможных варианта такого расположения и соответствующие каждому из них условия устойчивости системы:

6.1. Реальная область пересечений принадлежит области D_{ω}^+ возрастания функции параметра

$$D_{\omega}^R \subset D_{\omega}^+. \quad (16)$$

Такой случай расположения D_{ω}^R имеет место при условии

$$\omega_{\max} < \omega_{e\min}.$$

Ввиду известных асимптотических свойств корневых годографов [5, 6] можно сформулировать следующее утверждение.

Утверждение 1. При выполнении условия (16) семейство Z начальных точек годографов корневого портрета динамической системы, описываемой характеристическим уравнением (1), располагается в левой полуплоскости L корней:

$$Z \subset L. \quad (17)$$

Рассмотрим множество S участков s_i ветвей корневого портрета P системы:

$$S = \{s_i = [0, a_4(\omega_i)], i = 1, 2, \dots\}, \quad (18)$$

где $a_4(\omega_i)$ – значения функции (5) в точках на оси $i\omega$ с координатами ω_i ; $S \subset P$ и $S \subset L$ (см. (17)).

Тогда на основе (16) и (17) запишем

$$\bigcap_{i=1}^{\infty} s_i = \inf S = [0, \underline{a}_4(\omega_{\min})], \quad (19)$$

где $\underline{a}_4(\omega_{\min})$ – минимальное значение функции (5) в точке с координатой ω_{\min} (см. (14)).

Следовательно,

$$\forall a_4 \in [\underline{a}_4, \bar{a}_4] [a_4 \in [0, \underline{a}_4(\omega_{\min})] \rightarrow a_4 \in S \ \& \ P \subset L]; \quad (20)$$

$$\forall a_4 \in [\underline{a}_4, \bar{a}_4] [a_4 \notin [0, \underline{a}_4(\omega_{\min})] \rightarrow a_4 \notin S \ \& \ P \not\subset L]. \quad (21)$$

На основании (18)–(21) сформулируем следующее утверждение.

Утверждение 2. Для обеспечения асимптотической устойчивости динамической системы, описываемой семейством (1) интервальных характеристических уравнений и удовлетворяющей выражению (16), должно выполняться условие

$$\bar{a}_4 < \underline{a}_4(\omega_{\min}). \quad (22)$$

Определение 4. Доминирующими полиномами в семействе полиномов (1) назовем такие полиномы с постоянными коэффициентами, принадлежащие данному семейству, устойчивость которых гарантирует устойчивость всего семейства в целом.

Из условия (22) следует, что асимптотическая устойчивость системы определяется значением $\underline{a}_4(\omega_{\min})$. Это означает, что доминирующим в данном случае будет полином семейства (1), определяемый в точке с координатой ω_{\min} и значением функции параметра, равным $\underline{a}_4(\omega_{\min})$. В свою очередь точка с координатой ω_{\min} определяется выражениями (14), а именно коэффициентами $\bar{a}_1, \underline{a}_3$ полинома (1); минимальное значение параметра $\underline{a}_4(\omega_{\min})$ в данной точке определяется выражением (11), т. е. коэффициентом \underline{a}_2 .

На основании вышесказанного сформулируем следующее условие устойчивости.

Условие устойчивости 1. Если корневой портрет P (6) интервальной динамической системы, описываемой семейством характеристических полиномов (1), удовлетворяет выражению (16), то асимптотическая устойчивость такой системы обеспечивается, когда устойчивым является следующий полином данного семейства:

$$s^4 + \bar{a}_1 s^3 + \underline{a}_2 s^2 + \underline{a}_3 s + \bar{a}_4 = 0. \quad (23)$$

Условие устойчивости 1 используется для проверки системы на устойчивость, выражение (22) может быть использовано для параметрического синтеза систем.

6.2. Реальная область пересечений принадлежит области D_{ω}^- убывания функции параметра

$$D_{\omega}^R \subset D_{\omega}^-. \quad (24)$$

Такой случай расположения D_{ω}^R имеет место при условии

$$\omega_{\min} \geq \omega_{e_{\max}}.$$

С учетом известных асимптотических свойств корневых годографов [5, 6] можно сформулировать следующее утверждение.

Утверждение 3. При выполнении условия (24) семейство Z начальных точек годографов корневого портрета динамической системы, описываемой характеристическим полиномом (1), располагается в правой полуплоскости R корней:

$$Z \subset R. \quad (25)$$

Учитывая особенности асимптотических свойств годографов систем четвертого порядка [5, 6], можно заключить, что из выражения (25) следует неустойчивость всего семейства (1) в целом. В данном случае в семействе (1) не будет ни одного устойчивого полинома.

6.3. Реальная область пересечений полностью или частично принадлежит области D_ω^c чередования участков возрастания и убывания функции параметра:

$$D_\omega^R \subset D_\omega^c \vee D_\omega^R \cap D_\omega^c. \quad (26)$$

Такой вариант расположения D_ω^R имеет место, когда не выполняется ни одно из условий

$$\begin{aligned} \omega_{\max} &< \omega_{e_{\min}}, \\ \omega_{\min} &\geq \omega_{e_{\max}}. \end{aligned}$$

В данном случае определенная часть (P^+) годографов корневого портрета P пересекает границу устойчивости в областях возрастания и определенная (P^-) – в областях убывания функции параметра (5):

$$P = P^+ + P^-. \quad (27)$$

Возрастающая часть портрета (27), когда $P^- = \emptyset$, была рассмотрена в п. 6.1. Поэтому в п. 6.2 будем рассматривать только убывающую часть портрета P^- .

Рассмотрим семейство Z начальных точек годографов корневого портрета P^- . Учитывая особенности асимптотических свойств корневых годографов [5, 6], можно сформулировать следующее утверждение.

Утверждение 4. При выполнении условия (26) семейство Z начальных точек годографов корневого портрета динамической системы, описываемой характеристическим уравнением (1), может располагаться как в левой L , так и в правой R полуплоскости корней, т. е. могут быть следующие варианты расположения Z :

$$Z \subset L; \quad (28)$$

$$Z \subset (L + R); \quad (29)$$

$$Z \subset R. \quad (30)$$

Очевидно, что варианты (29) и (30) имеют место, когда

$$D_\omega^R \cap [z', \infty] \quad (31)$$

или

$$D_\omega^R \subset [z', \infty]. \quad (32)$$

Поскольку варианты (29)–(32) заведомо означают неустойчивость системы в целом, далее рассмотрим только вариант (28) расположения полюсов системы:

$$\omega_{\max} < \omega(z'), \quad (33)$$

где $\omega(z')$ – координата ω в точке z' (см. рис. 1).

Определим множество S участков s_i ветвей корневого портрета P системы выражением (18). Тогда на основе (26) и (28) запишем

$$\bigcap_{i=1}^{\infty} s_i = \inf S = [0, \underline{a}_4(\omega_{\max})], \quad (34)$$

где $\underline{a}_4(\omega_{\max})$ – минимальное значение функции (5) в точке с координатой ω_{\max} (14).

Следовательно,

$$\forall a_4 \in [\underline{a}_4, \overline{a}_4] [a_4 \in [0, \underline{a}_4(\omega_{\max})] \rightarrow a_4 \in S \ \& \ P \subset L]; \quad (35)$$

$$\forall a_4 \in [\underline{a}_4, \overline{a}_4] [a_4 \notin [0, \underline{a}_4(\omega_{\max})] \rightarrow a_4 \notin S \ \& \ P \not\subset L]. \quad (36)$$

На основе (26)–(36) сформулируем следующее утверждение.

Утверждение 5. Для обеспечения асимптотической устойчивости динамической системы, описываемой семейством (1) интервальных характеристических уравнений и удовлетворяющей выражению (26), должно выполняться условие

$$\overline{a}_4 < \min \{ \underline{a}_4(\omega_{\min}), \underline{a}_4(\omega_{\max}) \}. \quad (37)$$

Из условия (37) следует, что асимптотическая устойчивость системы для части P^- портрета (27) при выполнении условия (28) определяется значением $\underline{a}_4(\omega_{\max})$. Это означает, что доминирующим в данном случае будет полином семейства (1), определяемый в точке с координатой ω_{\max} и значением функции параметра, равным $\underline{a}_4(\omega_{\max})$. В свою очередь, точка с координатой ω_{\max} определяется выражениями (14), а именно коэффициентами $\underline{a}_1, \overline{a}_3$ полинома (1); минимальное значение параметра $\underline{a}_4(\omega_{\max})$ в данной точке определяется выражением (11), т. е. коэффициентом \underline{a}_2 . Поэтому для части P^- (27) корневого портрета системы с целью проверки асимптотической устойчивости достаточно проверить на устойчивость следующий полином семейства (1):

$$s^4 + \underline{a}_1 s^3 + \underline{a}_2 s^2 + \overline{a}_3 s + \overline{a}_4 = 0. \quad (38)$$

Учитывая, что портрет в данном случае является составным (27), т. е. одна его часть (P^+) пересекает границу устойчивости в области возрастания, а другая (P^-) – в области убывания функции параметра, для проверки устойчивости следует использовать полиномы (23) и (38). Сформулируем следующее условие устойчивости.

Условие устойчивости 2. Если корневой портрет P (27) интервальной динамической системы, описываемой семейством характеристических уравнений (1), удовлетворяет выражению (26), то асимптотическая устойчивость данной системы обеспечивается, когда устойчивыми являются следующие полиномы семейства:

$$s^4 + \overline{a}_1 s^3 + \underline{a}_2 s^2 + \underline{a}_3 s + \overline{a}_4 = 0; \quad (39)$$

$$s^4 + \underline{a}_1 s^3 + \underline{a}_2 s^2 + \overline{a}_3 s + \overline{a}_4 = 0. \quad (40)$$

Как показывают полученные выше результаты, в случае (26) для проверки асимптотической устойчивости системы также может быть использовано только одно уравнение с постоянными коэффициентами семейства (1). Какое именно, зависит от результатов проверки условия (37). Если в результате проверки окажется, что $\min \{ \underline{a}_4(\omega_{\min}), \underline{a}_4(\omega_{\max}) \} = \underline{a}_4(\omega_{\min})$, то проверку следует выполнять с использованием уравнения (39). Если же окажется, что $\min \{ \underline{a}_4(\omega_{\min}), \underline{a}_4(\omega_{\max}) \} = \underline{a}_4(\omega_{\max})$, то ее следует выполнять согласно (40).

Для определения значений коэффициентов (1), при которых выполняются соотношения (28) и (33), т. е. коэффициентов, определяющих подсемейство годографов корневого портрета, которые пересекают границу устойчивости в области $[0, z']$, используются формулы (5) и (14). Коэффициенты a_1 и a_3 при этом должны удовлетворять соотношению

$$\sqrt{\frac{\bar{a}_3}{a_1}} < \omega(z'), \quad \bar{a}_3 < a_1 \omega^2(z'). \quad (41)$$

Условия устойчивости 1 и 2 применяются для проверки системы на устойчивость, выражения (22) и (37) могут быть использованы для параметрического синтеза систем.

Пример. Коэффициенты исходного полинома (1) заданы в табл. 1.

Таблица 1

a_j	\underline{a}_j	\bar{a}_j
a_4	20	30
a_3	40	250
a_2	20	70
a_1	5	10
a_0	1	1

В табл. 2 приведены значения характеристик, полученные в результате вычислений по приведенным выше формулам.

Указанные в табл. 2 области показаны на рис. 2. Точки, соответствующие доминирующим полиномам (39), (40), обозначены соответственно r' и r'' .

Таблица 2

Значения характеристик	Тип области		
	Область экстремумов D_ω^e	Реальная область D_ω^R	$[z', z'']$
$\omega_{e\min}$	3,16	—	—
$\omega_{e\max}$	5,92	—	—
$a_{4e\min}$	100,04	—	—
$a_{4e\max}$	1223,96	—	—
ω_{\min}	—	2	—
ω_{\max}	—	7,1	—
$\underline{a}_4(\omega_{\min})$	—	64	—
$\underline{a}_4(\omega_{\max})$	—	-1532,97	—
$\omega(z')$	—	—	4,47
$\omega(z'')$	—	—	8,37

В рассматриваемом случае реальная область пересечений полностью охватывает область экстремумов: $D_\omega^e \subset D_\omega^R, [r', r''] \subseteq D_\omega^R$ (рис. 2). Очевидно, что данное семейство полиномов в целом является неустойчивым, так как его корневой портрет содержит подсемейство годографов, пересекающих границу устойчивости в области D_ω^- убывания функции параметра (см. п. 6.2 и рис. 2).

В области $Z_\omega = [z', z'']$ имеются полюсы, мигрировавшие в правую полуплоскость корней (см. (29)), что подтверждается отрицательным значением параметра $\underline{a}_4(\omega_{\max})$, приведенным в табл. 2.

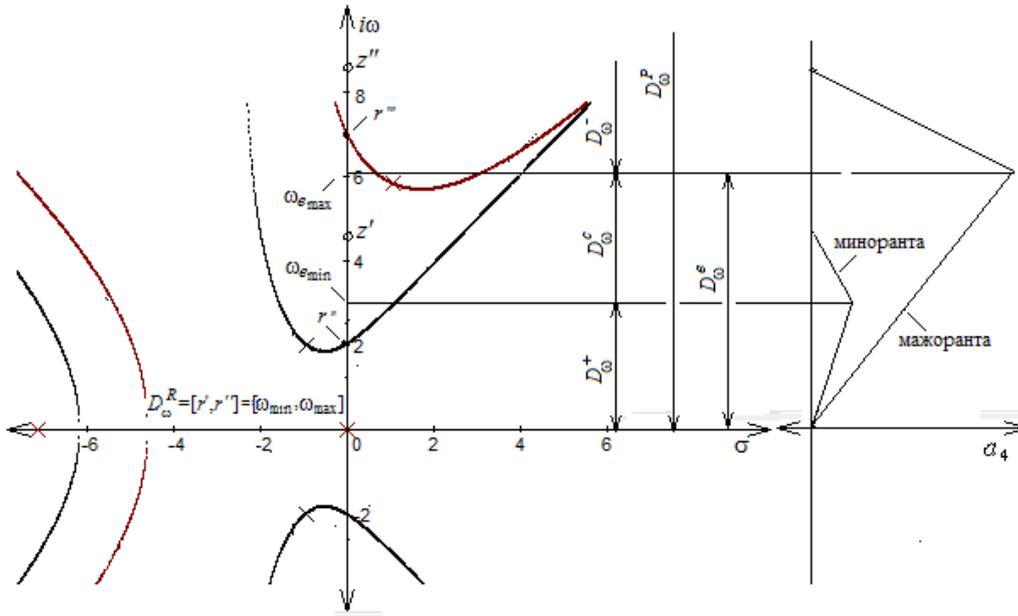


Рис. 2. Динамика корневого портрета интервальной динамической системы на границе асимптотической устойчивости

Доминирующие полиномы данного семейства имеют следующий вид:

$$s^4 + 10s^3 + 20s^2 + 40s + 30 = 0; \tag{42}$$

$$s^4 + 5s^3 + 20s^2 + 250s + 30 = 0. \tag{43}$$

Проверка полиномов на устойчивость показала, что полином (42), годограф которого пересекает границу устойчивости в точке $\underline{a}_4(\omega_{\min})$, является устойчивым, а полином (43), годограф которого пересекает границу устойчивости в точке $\underline{a}_4(\omega_{\max})$, имеет два корня, действительные части которых положительны, т. е. не является устойчивым. Это подтверждает сделанные выше выводы об устойчивости заданного семейства полиномов.

Выделим устойчивое подсемейство полиномов в заданном неустойчивом. Устойчивое подсемейство годографов, удовлетворяющее условиям (33) и (37), пересекает границу устойчивости в области, ограниченной интервалом $[r', z']$, поскольку в этом случае все начальные точки годографов семейства располагаются в левой полуплоскости корней (28) (см. разд. 4).

Для вычисления максимального значения коэффициента a_3 , которое определило бы устойчивое подсемейство годографов портрета, используем выражение (41):

$$\overline{a}_3 < \underline{a}_1 \cdot 4,47^2, \overline{a}_3 < 99,9. \tag{44}$$

На основании (44) определим $\overline{a}_3 = 80$.

После этого, используя условие (37), установим новые интервалы вариации a_4 , приняв

$$\omega_{\max} \in [r', z'], \omega_{\max} = 3. \tag{45}$$

Тогда

$$\underline{a}_4(\omega_{\min}) = 64, \underline{a}_4(\omega_{\max}) = 99.$$

На основании (37) примем $\overline{a}_4 < \underline{a}_4(\omega_{\min})$, $\overline{a}_4 = 60$ и запишем доминирующие полиномы для подсемейства:

$$s^4 + 10s^3 + 20s^2 + 40s + 60 = 0; \tag{46}$$

$$s^4 + 5s^3 + 20s^2 + 80s + 60 = 0. \tag{47}$$

Все корни полиномов (46) и (47) имеют отрицательные действительные части. Следовательно, согласно условию устойчивости 2 подсемейство годографов корневого портрета с измененными предельными значениями a_3 и a_4 ($\bar{a}_3 = 80, \bar{a}_4 = 60$) является устойчивым.

Таким образом, посредством установления верхней границы ω_{\max} реальной области пересечений согласно (45) и соответствующей настройки верхнего предела интервала изменения коэффициента a_3 в заданном табл. 1 неустойчивом семействе полиномов выделено подсемейство, которое является устойчивым.

На основе описанного выше метода в среде Visual Studio разработана программа, позволяющая выполнять исследование динамики систем с интервальной неопределенностью при помощи диаграммы распределения функции параметра вдоль границы устойчивости. На рис. 3 в качестве примера приведено главное окно данной программы, в котором показаны годографы корневого портрета для доминирующих полиномов системы, ограничивающие реальную область пересечений, и соответствующая диаграмма распределения функции параметра (5) вдоль границы устойчивости.

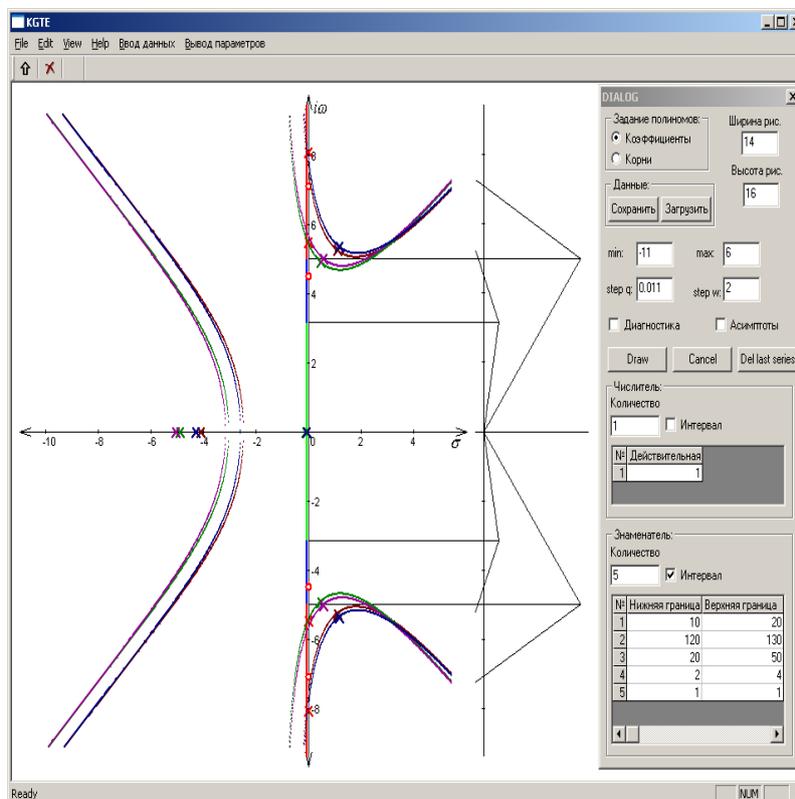


Рис. 3. Главное окно программы исследования динамических свойств интервальных систем на основе диаграммы распределения функции параметра вдоль границы устойчивости

В качестве входных данных программы вводятся коэффициенты или корни числителя и знаменателя передаточной функции системы. В поля «min», «max» и «step q» вводятся соответственно диапазон и шаг изменения значений δ , в поле «step w» – шаг изменения степени ω в уравнении корневого годографа. Полиномы задаются интервалами вариации коэффициентов. Предусмотрена возможность вычерчивания асимптот годографов.

Рассмотренный метод может получить свое развитие в ряде направлений, одним из которых является его обобщение на системы более высоких порядков.

Заключение

В работе выполнено исследование поведения динамических систем четвертого порядка в условиях интервальной вариации параметров с использованием корневых портретов. Установ-

лены закономерности поведения корневых портретов систем четвертого порядка с интервальной неопределенностью на границе асимптотической устойчивости. На основании этого сформулированы условия устойчивости и определены выражения для вычисления интервалов параметров, при которых обеспечивается устойчивость системы. В отличие от результатов В.Л. Харитонова [3] и Б.Д.О. Андерсона (B.D.O. Anderson) [4], которые предлагают для анализа устойчивости интервального семейства систем использовать соответственно четыре [3] и два (для систем четвертого порядка) [4] характеристических полинома этого семейства с постоянными коэффициентами, в настоящей работе благодаря применению корневых портретов систем установлено, что для анализа устойчивости подобных семейств систем может использоваться только один такой полином. Кроме того, применение корневого подхода позволяет не только использовать полученные в работе результаты в направлении анализа устойчивости [3, 4], но и выполнять на их основе робастный синтез динамических систем с параметрической неопределенностью.

Список литературы

1. Римский, Г.В. Автоматизация исследований динамических систем / Г.В. Римский, В.В. Таборовец. – Минск : Наука и техника, 1978. – 336 с.
2. Rimsky, G.V. Root locus methods for robust control systems stability and quality investigations / G.V. Rimsky, A.A. Nesenчук // 13th Triennial World Congress of International Federation of Automatic Control (IFAC) : Proceedings of the World Congress, San Francisco, USA, June 30 – July 05, 1996. – Oxford : Elsevier Science, 1996. – Vol. G. – P. 469–471.
3. Харитонов, В.Л. Об асимптотической устойчивости положения равновесия семейства систем линейных дифференциальных уравнений / В.Л. Харитонов // Дифференциальные уравнения. – 1978. – Т. XIV, № 11. – С. 2086–2088.
4. Anderson, B.D.O. On Robust Hurwitz Polynomials / B.D.O. Anderson, E.I. Jury, M. Mansour // IEEE Transactions on Automatic Control. – 1987. – Vol. AC 32, no. 10. – P. 909–913.
5. Дорф, Р. Современные системы управления / Р. Дорф, Р. Бишоп. – М. : Лаборатория базовых знаний, 2009. – 832 с.
6. Несенчук, А.А. Анализ и синтез робастных динамических систем на основе корневого подхода / А.А. Несенчук. – Минск : ОИПИ НАН Беларуси, 2005. – 234 с.
7. Поляк, Б.Т. Робастная устойчивость и управление / Б.Т. Поляк, П.С. Щербаков. – М. : Наука, 2002. – 303 с.
8. Поляк Б.Т. Управление линейными системами при внешних возмущениях / Б.Т. Поляк, М.В. Хлебников, П.С. Щербаков. – М. : ЛЕНАНД, 2014. – 560 с.

Поступила 21.10.2016

*Объединенный институт проблем информатики НАН Беларуси,
Минск, Сурганова, 6
e-mail: anes@newman.bas-net.by*

A.A. Nesenчук

INVESTIGATION OF THE INTERVAL DYNAMIC SYSTEMS BEHAVIOR AND SYNTHESIS OF THEIR CHARACTERISTIC POLYNOMIALS ON THE BASIS OF THE SYSTEM ROOT LOCUS PORTRAITS PARAMETER FUNCTION APPLICATION

Investigation of the 4th power dynamic systems behavior in conditions of the interval parameter variations is carried out on the basis of root locus portraits. Investigation of the system dynamics is carried out and the behavior regularities at the system stability bound are determined for root locus portraits of the 4th power systems with interval parametric uncertainties. On this basis the stability conditions are derived and expressions are determined for calculating intervals of parameter variation ensuring the system robust stability.

УДК 004.942:519.711.2:62.50

В.В. Ткаченко

ПРИМЕНЕНИЕ МЕТОДОВ ОБЩЕЙ ТЕОРИИ КОРНЕВЫХ ТРАЕКТОРИЙ В ИССЛЕДОВАНИИ СИСТЕМ МНОГОСВЯЗНОГО РЕГУЛИРОВАНИЯ

В рамках общей теории корневых траекторий рассматривается машинно ориентированный метод исследования систем управления с варьируемыми параметрами, которые входят в коэффициенты характеристического уравнения системы в виде сомножителей с произвольными показателями степени. Возможности метода демонстрируются на примере построения проекций областей устойчивости в плоскости варьируемых параметров двухканальной импульсной системы с перекрестными связями.

Введение

Система уравнений, описывающая многомерную систему управления с линейными стационарными звеньями или систему многосвязного регулирования (СМР), в матричной форме имеет вид [1]

$$C(p)y = B(p)u + D(p)f, \quad (1)$$

где y , u , f – матрицы-столбцы соответственно для векторов выходных величин, задающих и возмущающих воздействий; $C(p)$, $B(p)$, $D(p)$ – постоянные вещественные матрицы размерностью $n \times n$; p – оператор дифференцирования.

Основными задачами при исследовании систем управления являются проверки условий устойчивости и соответствия показателей качества заданным требованиям к переходным процессам. Для асимптотической устойчивости линейной системы по А.М. Ляпунову необходимо и достаточно, чтобы все корни ее характеристического уравнения были расположены слева от мнимой оси на комплексной плоскости корней. Область расположения корней в левой части полуплоскости, ограниченной линиями заданной степени устойчивости и заданной колебательности, определяет качество системы, оцениваемое временем регулирования и перерегулированием переходного процесса [2]. Условия устойчивости СМР формулируются аналогично случаю системы с одной регулируемой величиной и анализируются с использованием характеристического уравнения обобщенной передаточной функции СМР (1). При отсутствии возмущающих воздействий обобщенный характеристический полином СМР записывается как определитель матрицы $C(p)$ [2].

Характеристическое уравнение СМР в общем случае может содержать функциональные и связанные свободные параметры. В частности, при наличии перекрестных связей между каналами двухмерной системы ее характеристическое уравнение содержит варьируемый квадратичный параметр. В известных работах к такому случаю сводятся задачи, в которых возможно установление функциональной зависимости между варьируемыми параметрами, и задачи, в которых контуры основных каналов управления СМР имеют одинаковые передаточные функции [2–6].

1. Общее решение задачи нахождения границ области устойчивости в плоскости варьируемых параметров

В общей теории корневых траекторий Г.В. Римским [2, 7] обоснован принцип функциональной параметризации и разработаны методы для исследования устойчивости систем управления, которые дают возможность находить уравнения функциональных корневых траекторий, образуемых с помощью многозначных функций отображения для уравнений вида

$$\sum_{i=0}^n A_i(p)K^{n-i} = 0, \quad (2)$$

где K – действительная величина варьируемого параметра; $A_i(p) = \sum_{r=0}^l a_{ir} p^{l-r}$ – полиномы с постоянными коэффициентами a_{ir} [7].

Методы, разработанные в рамках общей теории корневых траекторий, позволяют решать широкий круг задач: анализ устойчивости линейных стационарных и нестационарных систем, параметрический синтез линейных и нелинейных многомерных систем, исследование импульсных систем, а также систем с запаздыванием и с распределенными параметрами. На основе этих методов разработаны пакеты прикладных программ [7], использование которых позволяет автоматизировать процесс аналитических преобразований и численных расчетов, сохраняя при этом приемлемую точность и обеспечивая наглядность для проектировщика благодаря графическому представлению результатов расчета. В частности, задача исследования системы с произвольным числом свободных параметров может быть сведена к построению проекций многомерной области устойчивости на плоскости двух параметров, т. е. к построению диаграмм Вышнеградского.

Рассмотрим задачу построения области устойчивости СМР в пространстве ее варьируемых параметров.

Пусть СМР имеет характеристическое уравнение вида

$$f(p) = \sum_{k=0}^m C_2^{m-k} \left(\sum_{i=0}^n C_1^{n-i} A_{ik}(p) \right) = 0, \quad (3)$$

где C_1 и C_2 – варьируемые параметры; $A_{ik}(p)$ – полиномы произвольного порядка с постоянными коэффициентами a_{ikr} от комплексной переменной $p = \delta + j\omega$.

Принцип последовательных преобразований плоскостей свободных параметров определяет прямую процедуру решения задачи построения заданных областей в пространстве коэффициентов алгебраического уравнения и пространстве его корней [2]. В общей теории корневых траекторий доказано, что результантные уравнения корневого годографа Теодорчика – Эванса, построенные с помощью уравнений годографа и его параметра, инвариантны относительно коэффициентных функций отображения. Методами понижения степени полиномов и построения результантных миноров достигаются упрощение и машинная реализация процедуры. Для вывода формул, по которым могут быть вычислены соответствующие заданным положениям полюсов на комплексной плоскости значения параметров C_1 и C_2 , построим модифицированные результанты. Элементами результантов являются полиномы, составленные из мнимых и действительных частей исходного полинома $f(p)$ [2, 7]. Результанты приравняем к нулю:

$$\text{Rez}_{C_2} f(p) = \begin{vmatrix} \sum_{i=0}^n C_1^{n-i} \text{Im}A_{i0}(p) & \sum_{i=0}^n C_1^{n-i} \text{Im}A_{i1}(p) & \cdots & \sum_{i=0}^n C_1^{n-i} \text{Im}A_{im}(p) & \cdots & 0 & \cdots \\ \sum_{i=0}^n C_1^{n-i} \text{Re}A_{i0}(p) & \sum_{i=0}^n C_1^{n-i} \text{Re}A_{i1}(p) & \cdots & \sum_{i=0}^n C_1^{n-i} \text{Re}A_{im}(p) & \cdots & 0 & \cdots \\ 0 & \sum_{i=0}^n C_1^{n-i} \text{Im}A_{i0}(p) & \sum_{i=0}^n C_1^{n-i} \text{Im}A_{i1}(p) & \cdots & 0 & \cdots \\ 0 & \sum_{i=0}^n C_1^{n-i} \text{Re}A_{i0}(p) & \sum_{i=0}^n C_1^{n-i} \text{Re}A_{i1}(p) & \cdots & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & 0 & \cdots & \sum_{i=0}^n C_1^{n-i} \text{Im}A_{ik}(p) & \cdots & \sum_{i=0}^n C_1^{n-i} \text{Im}A_{im}(p) \\ \cdots & 0 & \cdots & \sum_{i=0}^n C_1^{n-i} \text{Re}A_{ik}(p) & \cdots & \sum_{i=0}^n C_1^{n-i} \text{Re}A_{im}(p) \end{vmatrix} = 0. \quad (4)$$

Аналогичный вид имеет результат $\text{Rez}_{C_1} f(p)$, но с размерностью $2n$. При задании фиксированного вещественного C_1 формула (4) определяет траектории корней одного из корневых годографов алгебраического уравнения (1), а при дискретном задании – семейство траекторий, называемое обобщенным корневым годографом.

При $n = m = 1$ точки корневого годографа принадлежат действительной оси и траектории, задаваемой уравнением (аргумент p для упрощения записи полиномов далее опущен)

$$\text{Rez}_{C_2} f = \begin{vmatrix} \text{Im}(A_{00} + C_1 A_{10}) & \text{Im}(A_{01}(p) + C_1 A_{11}) \\ \text{Re}(A_{00} + C_1 A_{10}) & \text{Re}(A_{01}(p) + C_1 A_{11}) \end{vmatrix} = 0.$$

Значения C_1 (C_2) в заданной точке годографа находятся как решения квадратного уравнения

$$\begin{aligned} & \text{Im}A_{00} \text{Re}A_{01} - \text{Im}A_{00} \text{Re}A_{01} + (\text{Im}A_{00} \text{Re}A_{01} + \text{Im}A_{00} \text{Re}A_{01}) C_1 - \\ & - (\text{Im}A_{00} \text{Re}A_{01} + \text{Im}A_{00} \text{Re}A_{01}) C_1 + (\text{Im}A_{00} \text{Re}A_{01} - \text{Im}A_{00} \text{Re}A_{01}) C_1^2 = 0. \end{aligned} \quad (5)$$

2. Оценка устойчивости импульсной СМР на примере двухканальной системы стабилизации амплитуды и нейтрали вынужденных колебаний

В качестве примера рассмотрим двухканальную СМР для стабилизации амплитуды и нейтрали вынужденных колебаний $V(t) = V_m(t)\sin(2\pi t/T) + V_0(t)$ электромеханического вибратора вблизи его резонансной частоты. На привод виброэлемента этой системы (рис. 1) от генератора подается синусоидальный сигнал основной частоты ($1/T$). Амплитуда и среднее значение (постоянная составляющая) возбуждающего воздействия регулируются по результату сравнения временных интервалов реверса выходного сигнала, которые фиксируются с помощью концевых датчиков, расположенных вблизи крайних точек V_1 и V_2 линии колебаний.

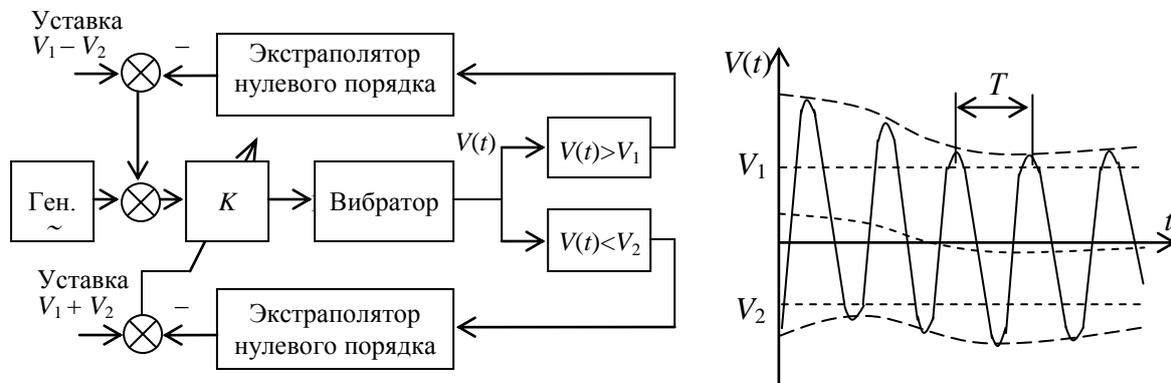


Рис. 1. Система стабилизации амплитуды и нейтрали вынужденных колебаний

Увеличение длительности интервалов реверса означает увеличение размаха колебаний в соответствующем полупериоде. Экстраполяторы, подключенные к концевым датчикам, преобразуют временные интервалы в напряжения, которые с соответствующими знаками сравниваются с уставками по амплитуде и нейтрали. Результаты сравнения подаются на входы умножителя и сумматора, через которые проходит сигнал возбуждения колебаний. В данном примере по сигналу первого из датчиков обеспечивается изменение амплитуды колебаний $V_m(t)$, а по сигналу второго – средней точки (нейтрали) колебаний $V_0(t)$. В практической реализации могут быть использованы другие варианты вычисления ошибок стабилизации амплитуды и нейтрали, в том числе интегральных и дифференциальных.

Рассматриваемый регулятор является системой с перекрестными связями на переменном токе (рис. 2). Как известно [1–3], в характеристических уравнениях таких систем возможно появление полюсов, не содержащихся по отдельности в каждом из основных контуров управле-

ния. Это накладывает дополнительные ограничения на выбор параметров для устойчивой работы системы. Кроме того, датчики регулятора, выполняя пороговые функции, придают системе существенно нелинейные свойства и усложняют ее анализ. Принимая во внимание, что фиксация текущих значений $V_M(t)$ и $V_0(t)$ осуществляется с определенным периодом вынуждаемых колебаний, процедуру линеаризации заменим переходом к рассмотрению двухсвязного регулятора в классе линейных импульсных систем с использованием дискретного преобразования Лапласа и подстановки $z = e^{pT}$.

Для оценки устойчивости режима стабилизации построим расчетную схему импульсной системы (рис. 2) и воспользуемся модифицированным z -преобразованием. Такое представление позволяет учесть тот факт, что такты квантования для положительных и отрицательных полувольт сдвинуты относительно друг друга на полупериод, а введение звена запаздывания $z^{-1} = e^{-pT}$ в один из каналов связи позволяет продемонстрировать возможность учета задержки вычисления ошибки рассогласования. Для оценки устойчивости в малом (при малых возмущающих воздействиях в режиме стабилизации) нелинейное преобразование интервалов реверса в каждой из полувольт колебаний в амплитуду и среднюю составляющую сигнала возбуждения можно представить в данном случае линейным звеном усиления с коэффициентами c_1 и c_2 . Эти коэффициенты являются свободными параметрами при определении области устойчивости.

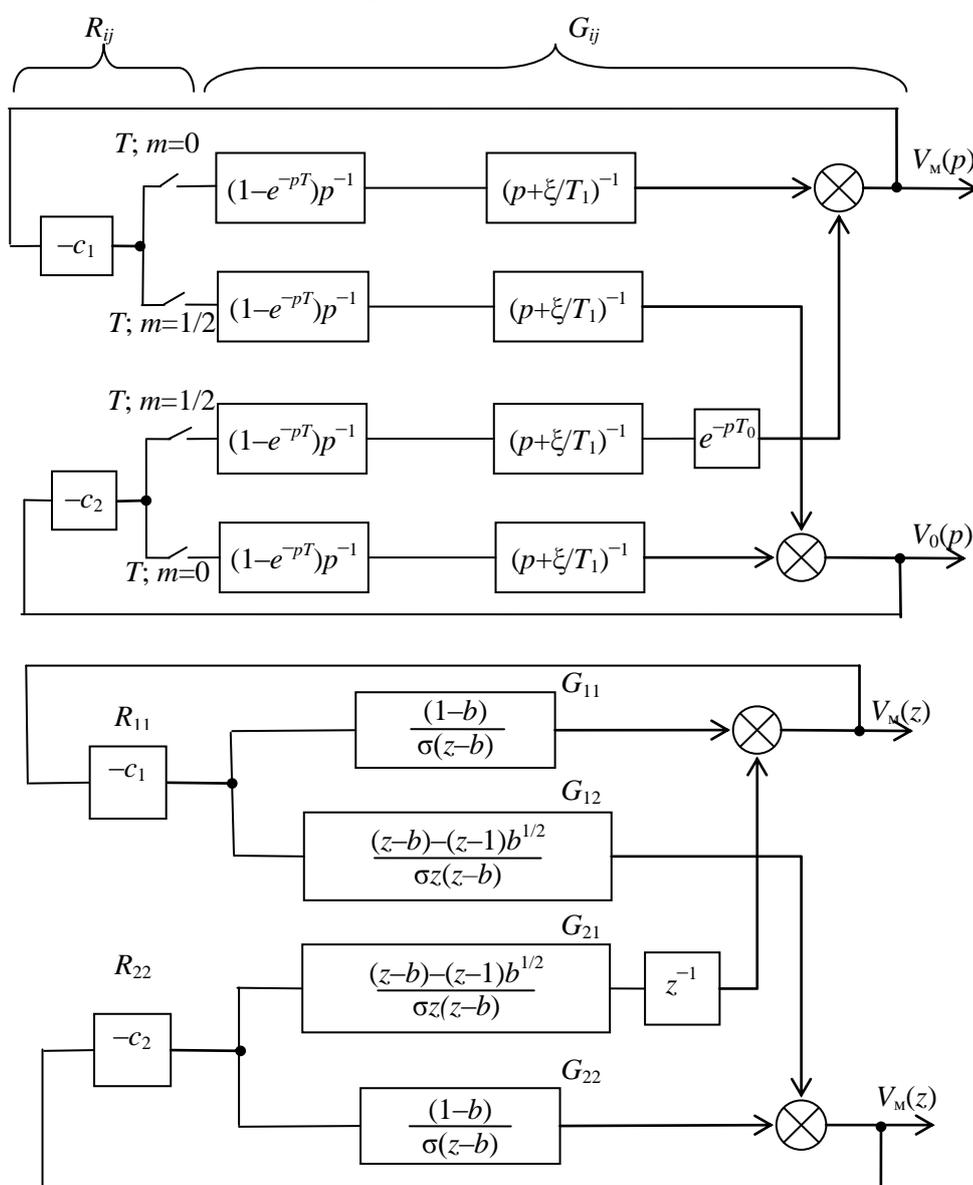


Рис. 2. Расчетная схема двухсвязной системы стабилизации колебаний

Для объекта, представляющего собой колебательное звено, его передаточные функции модулирующих сигналов на несущей частоте вблизи частотного резонанса в каждом из каналов СМР можно записать в виде передаточной функции аperiodического звена:

$$W(p) = K_0/T_1(p + \zeta/T_1), \quad (6)$$

где K_0 , T_1 , ζ – коэффициент усиления, постоянная времени и коэффициент затухания колебательного звена соответственно.

Используя формулы передаточных звеньев, приведенные на рис. 2, запишем характеристическое уравнение системы при отсутствии возмущающих воздействий:

$$[1 + G_{11}R_{11}][1 + G_{22}R_{22}] - G_{12}R_{11}G_{21}R_{22} = 0, \quad (7)$$

где G и R – коэффициенты связи для объекта и главного регулятора при $K_0/T_1 = 1$.

Подставляя вместо G и R передаточные функции звеньев, показанные на структурной схеме СМР в варианте ее дискретной модели (см. рис. 2), получим в левой части (7)

$$\left[1 + \frac{c_1(1-b)}{\sigma(z-b)}\right] \left[1 + \frac{c_2(1-b)}{\sigma(z-b)}\right] + \frac{c_1c_2[z-b-(z-1)\sqrt{b}]^2}{\sigma^2z^3(z-b)^2}, \quad (8)$$

где $b = \exp(\sigma T)$; $\sigma = \zeta/T_1$; T – период коммутации.

После приведения уравнения (7) к виду (3) получим

$$\Phi(z) + (c_1 + c_2)\Psi(z) + c_1c_2\Theta(z) = 0, \quad (9)$$

где Φ , Ψ , Θ – полиномы:

$$\begin{aligned} \Phi(z) &= \sigma^2z^3(z-b)^2; \Psi(z) = (1-b)\sigma z^3(z-b), \\ \Theta(z) &= z^3(1-b)^2 + (z + \sqrt{b})^2(1 - \sqrt{b})^2. \end{aligned} \quad (10)$$

Используя корневую методику, построим результаты

$$\operatorname{Re}z = \begin{bmatrix} \Phi_R + c_1\Psi_R & \Psi_R + c_1\Theta_R \\ \Phi_I + c_1\Psi_I & \Psi_I + c_1\Theta_I \end{bmatrix}, \quad \operatorname{Re}z = \begin{bmatrix} \Phi_R + c_2\Psi_R & \Psi_R + c_2\Theta_R \\ \Phi_I + c_2\Psi_I & \Psi_I + c_2\Theta_I \end{bmatrix}, \quad (11)$$

где индексы I и R соответствуют мнимым и действительным частям полиномов, и приравняем их нулю:

$$\begin{aligned} \Phi_I\Psi_R - \Phi_R\Psi_I + c_1(\Phi_I\Theta_R - \Phi_R\Theta_I) + c_1^2(\Psi_I\Theta_R - \Psi_R\Theta_I) &= 0; \\ \Phi_I\Psi_R - \Phi_R\Psi_I + c_2(\Phi_I\Theta_R - \Phi_R\Theta_I) + c_2^2(\Psi_I\Theta_R - \Psi_R\Theta_I) &= 0. \end{aligned} \quad (12)$$

Отметим, что благодаря совпадению передаточных функций главных каналов СМР частных полиномов в формулах (12) меньше, чем в (5). Для определения границ области устойчивости системы в пространстве параметров c_1 и c_2 перепишем характеристическое уравнение (9) в каноническом виде:

$$A(z) = \sum_{r=0}^l a_r z^{l-r}. \quad (13)$$

Согласно (10) $l = 5$. Коэффициенты a_r получаем суммированием коэффициентов из (9) и (10), стоящих перед z^{l-r} :

$$\begin{aligned} a_0 &= \sigma^2; \\ a_1 &= (1 - b)\sigma(c_1 + c_2) - 2b\sigma^2; \\ a_2 &= b^2\sigma^2 - (1 - b)b\sigma(c_1 + c_2) + (1 - b)^2c_1c_2; \\ a_3 &= (1 - b^{1/2})^2c_1c_2; \\ a_4 &= 2b^{1/2}(1 - b^{1/2})^2c_1c_2; \\ a_5 &= b(1 - b^{1/2})^2c_1c_2. \end{aligned} \quad (14)$$

Импульсная система устойчива, если корни ее характеристического уравнения на плоскости $z = \delta + j\omega$ лежат в пределах круга единичного радиуса. Прямому решению (13) совместно с условием $|z| = (\delta^2 + \omega^2) = 1$ сопутствуют вычислительные трудности. Пользуясь подстановкой $z^4 = z^2 = 1$, можно сразу понизить порядок (13):

$$z(a_0 + a_2 + a_4) + a_1 + a_3 + a_5 = zH + G = 0, \quad (15)$$

после чего сделать подстановку $z = (\omega^* + 1)/(\omega^* - 1)$, позволяющую с помощью билинейного преобразования отобразить круг единичного радиуса на левую половину плоскости $\omega^* = \omega_R^* + j\omega_I^*$, и воспользоваться результатным методом для определения границы области устойчивости в плоскости варьируемых коэффициентов.

После подстановки имеем:

$$(\omega^* + 1)H + (\omega^* - 1)G = (\omega^* + 1)(a_0 + a_2 + a_4) + (\omega^* - 1)(a_1 + a_3 + a_5) = 0, \quad (16)$$

Из равенства

$$\operatorname{Re} z = \begin{vmatrix} H+G & H-G \\ H+G & 0 \end{vmatrix} = H^2 - G^2 = 0$$

следуют уравнения $(a_0 + a_2 + a_4) = \pm(a_1 + a_3 + a_5)$. Подставляя в них выражения коэффициентов (14), получим уравнения кривых

$$\begin{aligned} \sigma^2 - ((1 - b)/(1 + b))\sigma(c_1 + c_2) + ((1 - b)^2/(1 + b)^2 + (1 - b^{1/2})^4/(1 + b)^2) c_1c_2 &= 0, \\ \sigma^2 - \sigma(c_1 + c_2) + 2c_1c_2 &= 0. \end{aligned} \quad (17)$$

Уравнения (17) связывают значения c_1 и c_2 , при которых некоторые корни полинома (13) располагаются на окружности единичного радиуса. Область устойчивости в плоскости свободных параметров определяется выбором пары значений варьируемых параметров в пределах каждой из ограниченных ветвями кривых (17) областей и проверкой принадлежности при этих значениях всех корней уравнения (14) внутренней части плоскости частот, ограниченной окружностью единичного радиуса. При установленной взаимосвязи коэффициентов a_r (14) для данного примера СМР таких областей не существует, иными словами, система неустойчива. Вывод подтверждается проверкой необходимых условий устойчивости дискретной системы по критерию Шура – Кона:

$$A(1) > 0; (-1)^l A(-1) > 0,$$

которые не выполняются для тех же диапазонов варьируемых параметров (13) b , c_1 и c_2 .

В практической работе выполнять аналитические преобразования (13)–(17) не требуется, так как необходимые графики получаются в результате численных решений уравнений (5). Коэффициенты таких уравнений определяются по численным расчетам из формул полиномов,

которые представляются в символической форме с применением метода многомерных коммутативных преобразований [7, 8]. Формулы (17) дают функциональное описание построенных границ в пространстве параметров, при переходе через которые изменение варьируемых параметров приводит к миграции корней через границу устойчивости. Соответствующие им кривые представляют собой гиперболы, повернутые так, что их оси направлены под углом 45° относительно координатных осей плоскости ($c_1; c_2$), а асимптоты каждой из гипербол смещены от координатных осей на некоторое расстояние. Порядок их получения демонстрирует необходимость выбора проектировщиком в каждом конкретном случае решаемой задачи последовательности действий с привлечением того набора методов, которыми располагает общая теория корневых траекторий и которые позволяют автоматизировать процесс создания их математических или компьютерных описаний. С помощью программных модулей осуществляются не только численные расчеты, но и предваряющие их необходимые аналитические преобразования над полиномами, что снимает проблемы достижения необходимой точности расчетов. Корневая методика позволяет также строить области заданного качества регулирования путем задания связи $f(\delta; \omega)$, например, при заданных требованиях к степени устойчивости и колебательности.

Заключение

С целью развития методов исследования многосвязных систем регулирования на основе общей теории корневых траекторий было рассмотрено построение плоских проекций областей их устойчивости в пространстве варьируемых параметров с использованием результата характеристического полинома анализируемой системы. Коэффициенты характеристического полинома могут содержать варьируемые параметры в виде сомножителей с произвольными показателями степени, что позволяет расширить круг задач для автоматизированного проектирования многоканальных или многомерных систем управления с перекрестными связями, таких как обеспечение устойчивости полета беспилотного аппарата или создание многосвязного робота с заданными требованиями по точности и времени выполнения движений.

Преимуществами рассматриваемого метода являются возможности:

- аналитического вывода уравнений проекций областей на плоскости свободных параметров, которые входят в характеристическое уравнение в виде сомножителей не только с первой, но и с любыми показателями степени (следовательно, возможности анализа многомерных систем с перекрестными связями);
- использования исходного уравнения в форме, не требующей приведения к каноническому виду, т. е. в той форме, в которой оно получается непосредственно из структурной схемы системы по передаточным функциям;
- вывода уравнений проекций машинным способом с использованием метода многомерных коммутативных преобразований [6, 7].

На примере двухсвязной системы с перекрестными связями, с помощью которой решается задача стабилизации амплитуды и нейтрали вынужденных колебаний объекта в зоне его резонансной частоты, продемонстрированы возможности применения корневого метода для перехода от исследования непрерывных систем к импульсным многосвязным системам с использованием операторного z -преобразования, понижения порядка анализируемого полинома и упрощения процедуры построения границ областей устойчивости в плоскости варьируемых коэффициентов.

Список литературы

1. Неймарк, Ю.И. Динамические системы и управляемые процессы / Ю.И. Неймарк. – М. : Кн. дом «ЛИБРОКОМ», 2009. – 336 с.
2. Римский, Г.В. Основы общей теории корневых траекторий систем автоматического управления / Г.В. Римский. – Минск : Наука и техника, 1972. – 328 с.
3. Gasparyan, O.N. Linear and Nonlinear Multivariable Feedback Control: A Classical Approach / O. N. Gasparyan. – England : John Wiley & Sons. Ltd, 2008. – 341 p.

4. Ильясов, Б.Г. Управление неустойчивыми объектами в составе многосвязной автоматической системы / Б.Г. Ильясов, Г.А. Сайтова, Е.А. Халикова // Современные проблемы науки и образования. – 2015. – № 1 (часть 2).

5. Ильясов, Б.Г. Алгоритм реконфигурации структуры многосвязной системы автоматического управления из условия устойчивости на основе частотных методов / Б.Г. Ильясов, Г.А. Сайтова, А.Ш. Назаров // Вестник УГАТУ. – 2012. – Т. 16, № 3(48). – С. 3–10.

6. Stengel, R. Singular Value Analysis of Linear Quadratic Systems / R. Stengel // Optimal Control and Estimation MAE 546 / Princeton University. – 2015. – 46 p. [Electronic resource]. – Mode of access : <http://www.princeton.edu/~stengel/MAE546Seminar14.pdf>. – Date of access : 04.09.2016.

7. Римский, Г.В. Автоматизация исследований динамических систем / Г.В. Римский, В.В. Таборовец. – Минск : Наука и техника, 1978. – 330 с.

8. Корневые методы исследования интервальных систем / под ред. Г.В. Римского. – Минск : Ин-т техн. кибернетики НАН Беларуси, 1999. – 186 с.

Поступила 31.10.2016

*Объединенный институт проблем
информатики НАН Беларуси,
Минск, Сурганова, 6
e-mail: tkach@newman.bas-net.by*

V.V. Tkachenko

**THE APPLICATION OF GENERAL THEORY OF ROOT
TRAJECTORIES METHODS IN RESEARCH OF MULTI-CHANNEL
CONTROL SYSTEM WITH CROSS-CONNECTIONS**

Computer aided research method on the basis of general theory of root trajectories is considered in the frame of study of multi-channel control systems with varying parameters included as cofactors with arbitrary exponents in coefficients of the characteristic equation of the system. Possibilities of the method are demonstrated on the example of constructing stability regions projections on the plane of varying parameters of dual-channel pulse system with cross-connections.

УДК 004.912

Г.Э. Романюк, М.С. Ходар

АНАЛИЗ ЛЕКСИКОГРАФИЧЕСКОЙ ОКРАШЕННОСТИ ХУДОЖЕСТВЕННЫХ ТЕКСТОВ

Предлагается программа анализа текста для нахождения и распознавания цветовых образов, используемых в тексте. С помощью программы фиксируются данные, строится матрица цветов и цветовые диаграммы. Полученные данные экспортируются в Excel для последующего изучения и анализа статистики цветовых образов и определения их зависимости от вида текста, жанра и автора.

Введение

Распознавание образов (объектов, сигналов, ситуаций, явлений или процессов) – едва ли не самая распространенная задача, которую человеку приходится решать практически ежедневно от первого до последнего дня своего существования.

Сотни исследований и тысячи экспериментов в области изучения филологии, семантики и психологии были проведены за последние годы [1]. Было доказано множество разнообразных гипотез [2, 3]. В настоящей работе исследуется содержание различных цветовых образов в тексте и их влияние на сознание человека.

Теоретической базой работы является обширный список печатных изданий и интернет-ресурсов, посвященных описанию компьютерных программ, проблем компьютерной лингвистики, результатов исследований в области структурного литературоведения, существующим теориям цвета [4].

Объем информации, доступной в сети Интернет, растет с каждым годом. При этом большая часть этой информации представляет собой тексты на естественном языке. Научной новизной результатов настоящей работы является возможность извлечения цветовой информации из таких текстов и ее применение в научно-исследовательской и практической области при изучении влияния цвета на человека.

1. Основные методы анализа текстов

В разделе дается общее описание процесса анализа текста и рассматриваются основные проблемы [5]. На вход любой системы для анализа текста, как правило, поступает текст в исходном формате, т. е. в популярных форматах документов DOC, OAT, RTF, PDF, HTML. Такие документы кроме самого текста содержат также форматирование, сноски, «лишние» участки текста (рекламу). Первой сложностью при анализе текста является его извлечение из документов, поданных на вход. Далее текст следует пропустить через графематический анализатор, задача которого – определить границы слов, предложений, абзацев, а также отметить числа, знаки препинания, прочие символы и последовательности символов, не являющиеся словами, но играющие важную роль в предложении.

Когда предложения размечены и слова локализованы, начинает работу морфологический анализатор. Применяя имеющийся словарь и набор словоизменительных правил, этот анализатор дополняет каждое слово набором грамматических характеристик. На этапе морфологического анализа также возникает ряд сложностей с обработкой опечаток, особенно если они допущены в окончании слова, играющем ключевое значение в определении набора грамматических характеристик. Ряд сложностей возникает при анализе неизвестных слов, а также заимствованных из других языков.

Следующий этап анализа – синтаксический. Наиболее популярным подходом к этому анализу является построение дерева синтаксического разбора предложения. В процессе построения дерева исходное предложение перестраивается в древовидную структуру, представляющую собой аналог дерева синтаксического разбора некоторой формальной грамматики. При частичном

синтаксическом анализе полное дерево разбора не строится, а вместо этого анализатор сосредоточивается на поиске заранее определенных синтаксических конструкций в тексте.

Такой подход вполне приемлем, так как естественный язык не имеет грамматики в привычном математикам смысле. В естественных языках, как правило, достаточно вольные правила построения грамматических конструкций, слабо поддающиеся формализации. В достаточной степени формализуемы лишь некоторые синтаксические отношения:

- согласования и зависимости при построении словосочетаний;
- валентности глаголов и отглагольных причастий;
- ссылки на контекст (различные типы местоимений);
- общие правила построения предложений (простых предложений, сложносочиненных, сложноподчиненных и других сложных предложений).

Заключительным этапом анализа текста является семантический анализ результатов, полученных на этапе синтаксического анализа. Для этого этапа не существует устоявшихся моделей и подходов. В большинстве систем роль семантического анализатора играет эвристически реализованный модуль, осуществляющий поставленную перед ним задачу. Ясно, что даже при наличии дерева полного синтаксического разбора текста все равно сложно реализовать алгоритм, извлекающий полезную информацию на основе полученного дерева [1].

Одним из результатов исследования является новый подход к анализу текста – потоки интерпретаций. Потоки интерпретаций представляют собой множество пар (участок текста) и некоторую информацию. При этом не ставится никаких ограничений на тип информации и ее формат. Таким образом, основное отличие предлагаемого подхода от традиционного состоит в том, что если традиционно принято делить процесс анализа на этапы, жестко разделяя порядок исполнения этапов анализа, а также результат каждого этапа, то в подходе, основанном на потоках интерпретаций, предлагается фиксировать лишь формат результатов анализа.

Основные методы и методики анализа текстов:

Интеннт-анализ – метод, позволяющий реконструировать интенции автора по его тексту, поскольку для выявления и квалификации интенций опора на отдельные слова и предложения малопродуктивна [1]. Экспертное выявление и идентификация речевых интенций предоставляют возможность очертить их круг в текстах разной тематики и направленности, т. е. охарактеризовать их качественно, поэтому исследовательская задача у использующих метод интеннт-анализа состоит в экспертном (т. е., по сути, субъективном) оценивании характера интенций, их размытости и неясности понимания. Метод состоит из последовательных этапов: выделения круга обсуждаемых тем и вопросов, определения связей между объектами, затем кодификаций дескрипторов. Далее проводится оценка групп объектов по нескольким интегральным измерениям, полученные значения усредняются и определяются интегральные значения каждого объекта по указанным параметрам. При исследовании текстов СМИ интеннт-анализ позволяет решать проблемы социально-психологического и общесоциального плана, например влияния средств массовой коммуникации на индивидуальное и групповое сознание.

Ethnograph – методика, предназначенная для качественного (содержательного) анализа данных интервью, фокус-групп, дневников и пр. [1].

Процесс работы с данными построен на основе того, что осуществляется поиск необходимых данных, отбор результатов поиска и анализ всего отобранного материала. Для каждого из звеньев этого процесса предусмотрен соответствующий набор процедур.

Leximancer – методика, направленная на выявление ключевых тем (key-themes) и концептов (concepts) в электронных документах. В отличие от других аналогичных методик в своей основе соединяет множество подходов, таких как вычислительная лингвистика, контент-анализ, информационные науки, физика, теория сетей и пр. Позволяет наглядно представлять результаты анализа и вычислений в виде карты концептов и разнообразных таблиц и диаграмм.

Minnesota Contextual Content Analysis (MCCA) – метод, позволяющий осуществлять контекстуальный анализ (анализ слов в пространстве четырех социальных контекстов: практики, традиций, эмоций, анализа) и анализ основных мыслей и образов, раскрытых в тексте. Для каждого вида анализа разработаны и стандартизированы нормы. Помимо указанных главных функций метод позволяет выводить статистику, проводить анализ слов, частотный анализ слов и категорий и т. д. [2].

Контент-анализ – самый распространенный метод, имеющий множество вариаций в различных методиках, позволяющий проводить качественно-количественный анализ содержания текстовых массивов с целью последующей интерпретации выявленных числовых закономерностей. Заключается в оценке частотного распределения слов, словосочетаний словоформ и других единиц анализа (число их вариаций теоретически безгранично) относительно текста. Результатом анализа является частота, относительный и удельный вес, вероятность встречаемости и пр., на основе чего делается качественный или количественный вывод в зависимости от выдвинутой гипотезы. Контент-анализ может быть проведен при помощи широкого спектра методик [1].

PROTAN – автоматизированная система контент-анализа, которая предназначена для анализа любого текста (рассказов, клинических интервью, научных публикаций, названий или резюме научных журналов, поэзии, рекламных материалов и др.). Ограничения PROTAN обусловлены статистическими ограничениями, отсутствием словарей, необходимых для того, чтобы анализировать специфический вид текста. Текст для анализа должен быть представлен в стандартной кодировке. Методика позволяет решать две основные задачи: определение структуры текста (при помощи семантических словарей), определение основных тем и идей текста (на основе информации, содержащейся во взаимосвязях между единицами анализа текста).

Yoshikoder – методика контент-анализа текста, включающая разработку и использование словарей, автоматический поиск по ключевым словам в контексте [4].

Фоносемантический анализ текста (или слова) – метод, основанный на оценке его звучания безотносительно содержания. Заключается в сопоставлении системы сочетаний фонем в конкретном тексте или слове с их стандартизированными оценками по ряду биполярных шкал. Результатом фоносемантического анализа является профиль выраженности оценочных шкал в стандартизированном семантическом пространстве, на основании которого делается заключение о возможном воздействующем эффекте текста на читателя. Однако ввиду крайней специфичности единиц анализа и непроработанного механизма соотнесения с содержанием текста (и отсутствия контроля факторов, влияющих на процесс осмысления текста) результаты этого метода представляются многим исследователям сомнительными и не обладающими внешней валидностью, что не отменяет адекватность применения данного метода для узкоспециализированных исследований.

Фоносемантический анализ реализован в следующих методиках и методах [5]:

Vaal – методика, в которой реализованы алгоритмы оценки фонетического воздействия на человека слов и текстов русского языка, причем в основе эмоционального воздействия фонетики слова и текста на подсознание человека лежат психофизиологические механизмы. Дает возможность анализировать готовые тексты с точки зрения такого воздействия, составлять новые с заданным вектором воздействия, выявлять личностно-психологические качества авторов текста, проводить углубленный контент-анализ и делать многое другое.

DIATON – методика экспертизы суггестивных текстов, основанная на фоносемантическом анализе и ориентированная на оценку скрытых особенностей, которые сложно осознать: фоносемантических, ритмических, структурных характеристик текста.

Дискурс-анализ, или дискурсивный анализ, – совокупность методик и техник интерпретации текстов или высказываний как продуктов речевой деятельности, осуществляемой в конкретных общественно-политических обстоятельствах и культурно-исторических условиях. Этот метод ориентирован прежде всего на изучение лингвистического уровня в структуре социальной коммуникации как доминирующего на протяжении определенного исторического периода развития общества и культуры. Сам метод заключается в последовательности ряда операций: фиксации изучаемого материала; выделении его формальных характеристик; обозначении контекста как коммуникативной ситуации; выборе направления и стратегии анализа; теоретическом дифференцировании и структурировании этапов исследования; определении техники и средств анализа при использовании конкретной модели исследования; дефиниции единиц анализа; проверке системы категорий в теории и на эмпирическом материале; осуществлении основных этапов исследования (описания, реконструкции, интерпретации); фиксации результатов исследования, их обобщении, истолковании и структурировании. Дискурс-анализ позволяет выделить не только существенные характеристики социальной коммуникации, но и второстепенные, содержательные и формальные показатели (например, тенденции в вари-

тивности речевых формул или построении высказываний). Дискурс-анализ широко применяется в социологических и политических исследованиях и отчасти реализован в таких программах, как САТРАС. Это методика анализа текста, написанного на любом языке. Она основана на системе Galileo, которая представляет собой комплекс теории и методов, направленных на научное изучение когнитивных и культурных процессов. САТРАС позволяет выявлять основные идеи текста без предварительного кодирования и лингвистического анализа.

Нарративный анализ – метод обобщения прошлого опыта при помощи соотнесения последовательности слов в предложении и последовательности реальных (как предполагается) событий. Позволяет осуществлять количественную оценку текста. В отличие от контент-анализа, который может быть применен к любым текстам, нарративный анализ ориентирован на особые тексты, содержащие рассказ. Преимуществом нарративного анализа по сравнению с кластерным является то, что оценка производится по конкретным категориям (субъект, действие, объект), а не по произвольно выбранным исследователем, исходя из его задач. В класс нарративных текстов входят разнообразные истории: от художественных и исторических текстов (мифов, легенд, летописей и пр.) до газетных статей, в которых описываются произошедшие события. Нарративный анализ используется совместно с другими методами анализа текста и реализован в известных методиках.

LIWC – методика, обладающая 68 встроенными словарями (лингвистические, психологические конструкты и др.), которые представляют собой пространство для оценки того, в какой степени испытуемые используют слова тех или иных категорий (например, позитивные и негативные слова). Методика позволяет осуществлять нарративный анализ текста, синтаксический анализ, интент-анализ и ряд других функций.

PC-ACE – методика, предназначенная для кодирования событий и позволяющая организовывать сложную текстовую информацию, хранить ее, осуществлять поиск необходимых данных, характеризующихся сложной структурой. Методика применима для всего спектра социальных наук и позволяет осуществлять качественный (содержательный) анализ данных.

Экспертная оценка текста – группа методов, в которую входят различные экспертизы текста, классификацию которых, согласно А.А. Леонтьеву [5], можно представить в следующем виде:

а) автороведческая экспертиза, направленная на установление автора текста или выявление категориальных признаков вероятного автора: пол, возраст, национальность, место рождения, место долговременного проживания, уровень образования и пр. ;

б) экспертиза, направленная на установление временных признаков автора текста (эмоциональное состояние и пр.);

в) экспертиза, направленная на установление тех или иных условий создания исследуемого текста (также экспертиза аутентичности записей при интервью);

г) экспертиза, направленная на установление преднамеренного искажения сведений, высказываемых в тексте;

д) экспертиза, направленная на установление определенных признаков (оскорбление, призыв и пр.).

Для осуществления данных экспертных оценок применяется комплекс методик: перефразирования текста или законченного фрагмента текста; семантического шкалирования, например методика семантического интеграла (В.И. Батов, Ю.А. Сорокин); свободного ассоциативного эксперимента; предикативного анализа текста. Существуют также компьютеризированные варианты данных экспертиз.

Графематический анализ – метод, создающий базу для последующего морфологического и синтаксического анализа на основе выделения слов, цифровых комплексов, формул и т. д. Анализ направлен на разбивку текста на слова, разделители и т. д.; сборку слов, написанных в разрядку; выделение устойчивых оборотов, фамилии, имени, отчества, даты и т. п.; выделение электронных адресов и имен файлов; выделение предложений из входного текста абзацев, заголовков, примечаний.

Морфологический анализ направлен на определение множества морфологических интерпретаций каждого из слов текста, состоящего из таких параметров, как лемма, морфологическая часть речи, набор общих граммем, множество наборов граммем. Морфологический анализ

реализован в большинстве методик, так как является основой для других видов анализа текста. В качестве примера можно отметить реализацию рассматриваемого метода.

ATLAS.ti – методика, позволяющая анализировать большие объемы текста, разнообразные графики, аудио- и видеоинформацию. Может применяться в социальных и экономических науках, маркетинге и менеджменте, теологии. Методика рассчитана на проведение качественного (содержательного) анализа данных и включает средства исследования текста, управления текстом, сравнения и пр.

Textanz – методика, предназначенная для частотного анализа текста на уровне слов, фраз, словоформ. С помощью данной методики можно осуществлять анализ любого текста. Она также применима для синтаксического анализа.

Цветонаименования в художественном тексте служат изобразительным средством. Классическая литература – достоверный источник сведений для понимания особенности цветов как эмоционально-перцептивных эталонов. Цвет в литературе – средство выражения, экспрессии. Он не только служит для моделирования абстрактного зрительного образа, но и сам встречается в определенных экспрессивных контекстах и жизненных ситуациях, которые определенным образом окрашены эмоционально и составляют органическую часть идейного содержания художественного произведения. Можно реконструировать внутреннюю форму цветов как символов эмоциональных состояний, изучая их употребление в произведении. Средством для такой реконструкции служат контекстные ситуации, в которые писатель вводит цвет как изобразительное средство.

2. Постановка задачи

Распознавание представляет собой задачу преобразования входной информации, в качестве которой уместно рассматривать некоторые параметры, признаки распознаваемых образов, в выходную, представляющую собой заключение о том, к какому классу относится распознаваемый образ.

В данном исследовании ставятся следующие задачи:

- разработать алгоритм и концепцию анализа текста на наличие цветовых образов;
- создать программу, которая сможет анализировать текст на содержание цветовых образов;
- научить программу отображать цветовую матрицу текста;
- разработать функцию выведения численной и графической статистики цвета в тексте;
- разработать методы применения программы для нахождения зависимостей между смыслом текста, его цветовым содержанием и влиянием на человека.

Успешность решения поставленных задач основывалась на наличии программно-технического обеспечения, определенной теоретической базы, перспективных идей и практических разработок.

3. Концептуальная модель разрабатываемой программы Text Analyst

Цель информационного моделирования – создание концептуальной схемы предметной области. Эта схема (или просто модель) в упрощенном виде отражает наиболее важные для пользователей информационные объекты предметной области и связи между ними (рис. 1).

В разрабатываемом приложении пользователь является ключевым объектом системы, который управляет и инициирует процессы программы. Он должен самостоятельно выбирать и загружать исследуемый текст в программу. Именно пользователь решает, что именно следует проанализировать в тексте, и активирует те или иные функции программы. Все функции связаны с пользователем ассоциативной связью. Между двумя классами (объектами) существуют разные типы отношений. Самым базовым типом отношений является ассоциация (association). Это означает, что два класса связаны между собой. Обычно такое отношение используется на ранних этапах дизайна, чтобы показать, что зависимость между классами существует. Только ассоциативных связей недостаточно, чтобы на диаграмме классов показать все связи между объектами.

Классовые объекты «анализируемый текст», «цветовая диаграмма», «цветовая палитра», «матрица цветов» и Excel-файл связаны между собой такими связями, как агрегация и композиция. Обе они моделируют отношение «является частью» (HAS-A Relationship) и обычно выражаются в том, что класс целого содержит поля (или свойства) своих составных частей. Грань между ними достаточно тонкая, но важная (особенно в контексте управления зависимостями). Чтобы легче запомнить визуальную нотацию, ромбик всегда находится со стороны целого, а простая линия – со стороны составной части; закрашенный ромб означает более сильную связь – композицию, незакрашенный ромб показывает более слабую связь – агрегацию.

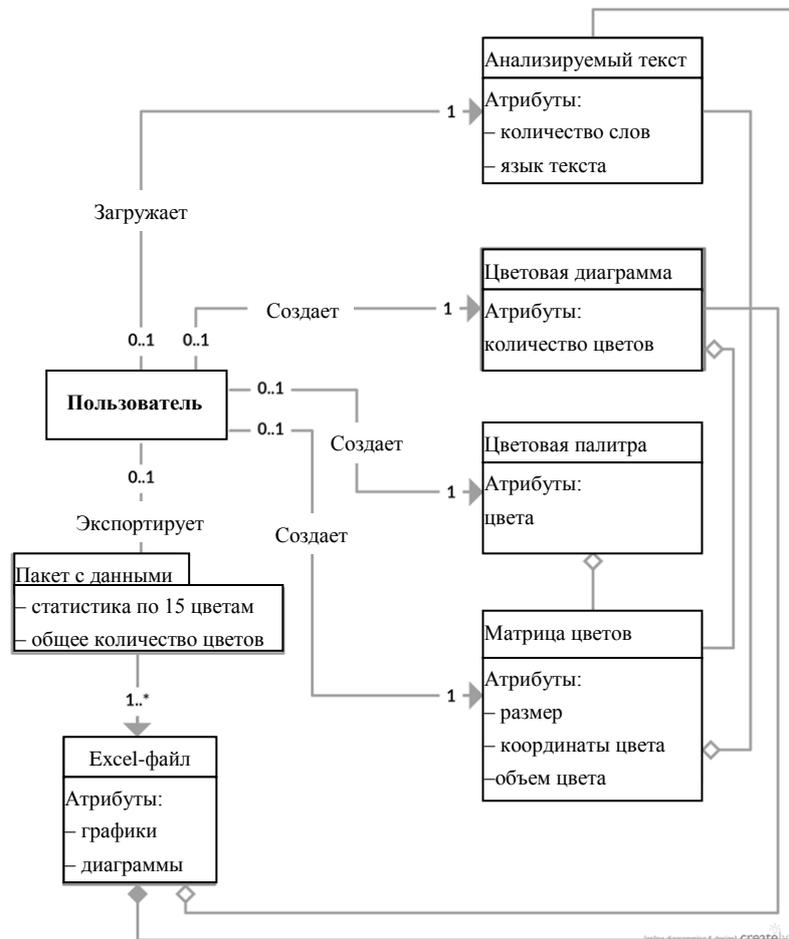


Рис. 1. Диаграмма классов

Разница между композицией и агрегацией заключается в том, что в случае композиции целое явно контролирует время жизни своей составной части (часть не существует без целого), а в случае агрегации не контролирует (например, составная часть передается через параметры конструктора):

```

class CompositeCustomService
{
//Композиция
    private readonly CustomRepository _repository
    = new CustomRepository();
    public void DoSomething()
    {
//Используем _repository
    }
}
  
```

```

class AggregatedCustomService
{
//Агрегация
    private readonly AbstractRepository repository;
public AggregatedCustomService(AbstractRepository repository)
{
    _repository = repository;
}
public void DoSomething()
{
//Используем repository
}
}

```

CompositeCustomService для управления своими составными частями использует композицию, а AggregatedCustomService – агрегацию. При этом явный контроль времени жизни обычно приводит к более высокой связанности между целым и частью, поскольку используется конкретный тип, тесно связывающий участников между собой. С одной стороны, такая жесткая связь может не являться чем-то плохим, особенно когда зависимость стабильна. С другой стороны, можно использовать композицию и контролировать время жизни объекта без привязки к конкретным типам.

Концептуальная модель программы анализа текста в виде диаграммы классов помогает разработчикам на ранней стадии понять, какие объекты будут находиться в системе и какие классы следует создать при написании кода, что значительно сокращает время разработчика на понимание системы в целом. Существует несколько достаточно объективных критериев для определения связности дизайна по диаграмме классов: большие иерархии наследования (глубокие или широкие иерархии) и повсеместное использование композиции, а не агрегации говорят о сильно связанном дизайне.

Ключевой функцией разрабатываемого приложения является сравнительный анализ различных текстов с целью нахождения зависимостей и связей среди распознанных цветовых образов. Программа выполняет ряд основных и несколько второстепенных функций (рис. 2).



Рис. 2. Диаграмма вариантов использования

Основными функциями приложения Text Analyst являются поиск цветов, подсчет их количества, вычисление объема, построение цветовой матрицы текста на основании этих данных и возможность статистического анализа полученных данных в программе Excel.

Второстепенными, вспомогательными функциями анализа являются построение цветowych диаграмм, цветовой палитры, поиск и подсчет цветов. Данные функции второстепенны, так как они не участвуют в решении задач, которые поставлены в настоящей работе, но они служат вспомогательным средством для визуализации, работы с данными и проверки полученных данных. Поэтому даже с учетом того, что эти дополнительные функции не являются обязательными, они значительно расширяют функциональность приложения для исследования текстов.

Функциональная спецификация требований к программному средству основана на диаграмме вариантов использования (см. рис. 2). Разрабатываемое программное средство распознавания цветовых образов на основе анализа текстов должно обеспечивать выполнение следующих функций:

- работу с пользовательским интерфейсом;
- анализ текстов;
- построение диаграмм;
- построение цветовой матрицы;
- импорт данных в Excel.

4. Описание программы Text Analyst

Пользовательский интерфейс программы Text Analyst (рис. 3 и 4) предназначен для анализа текстов на основе цветовых образов.

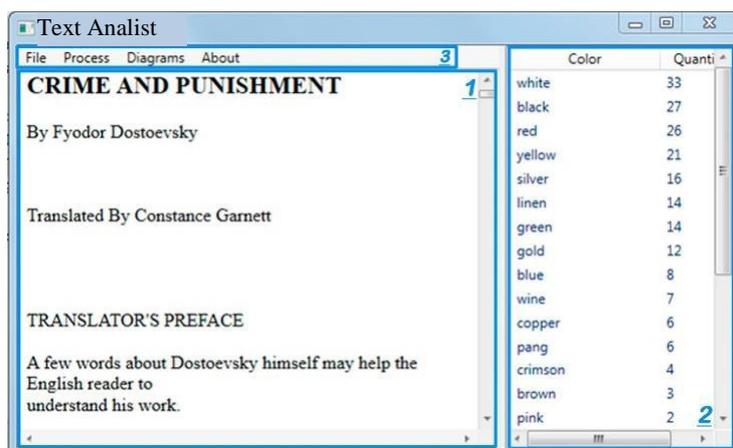


Рис. 3. Пользовательский интерфейс программы

Программа состоит из следующих основных блоков:

- 1 – текстовое поле, в котором содержится анализируемый текст;
- 2 – панель цветов, которая отображает, в каком количестве и какие цвета были найдены в тексте;
- 3 – панель инструментов, содержащая набор функций для анализа текста.

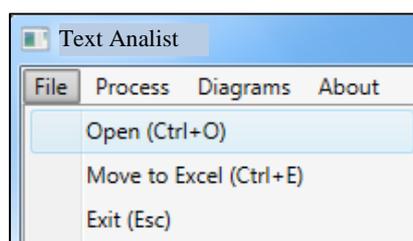


Рис. 4. Главное меню программы

Нажав на кнопку **File -> Open**, пользователь может выбрать нужный текстовый файл формата txt или RTF. Только после загрузки текста можно приступить к его анализу.

Для того чтобы выполнить поиск цветов в тексте, на панели инструментов необходимо нажать **Process -> Find Colors**. После этого появляется список цветов и их количество во втором блоке. Нажав на имя колонки, можно отсортировать данные по возрастанию или убыванию как по названиям цветов, так и по количеству их упоминаний (рис. 5).

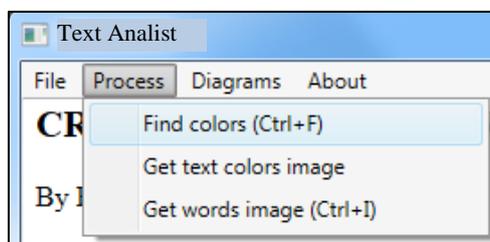


Рис. 5. Меню поиска текста и построения изображений

При нажатии **Process -> Get words image** программа анализирует текст и строит его цветовую матрицу. На примере (рис. 6) загружена книга Ф.М. Достоевского «Преступление и наказание». Видно, что чаще всего в книге упоминается белый цвет. Далее идут черный, красный, желтый и серебряный по частоте упоминания. Программа также не упускает такие цвета, как медный, цвет вина, малиновый закат, розовый, кремовый, горчичный, кукурузный, лимонный, лавандовый и т. д.

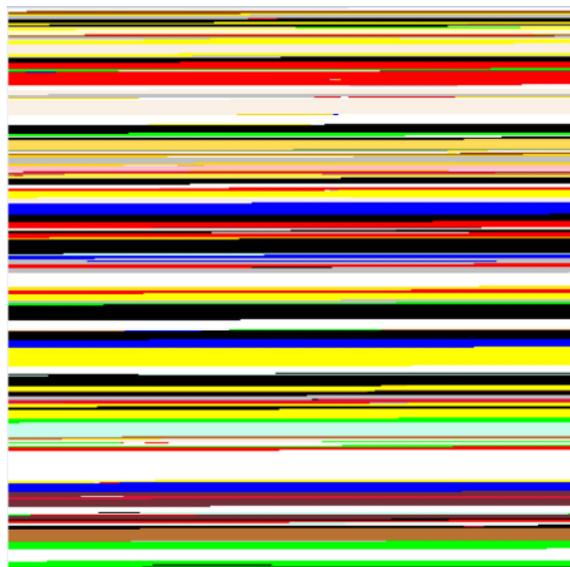


Рис. 6. Цветовая матрица романа Ф.М. Достоевского «Преступление и наказание»

На вход в программу подается текст в формате txt или RTF. Первым делом программа считает количество слов в тексте по количеству пробелов. Количество слов принимается за величину L . Извлекая из этого числа корень квадратный и округляя полученную величину до следующего целого числа, программа получает размер квадратной таблицы, которая заполняется цветами.

Приведем пример входного текста (рис. 7). Всего в тестовом тексте 100 слов. Это значит, что размер будущей таблицы будет $L = \sqrt{100} = 10$. Если количество слов в тексте, к примеру, 28 567, то размер таблицы будет $L = \sqrt{28567} \approx 169,01$. Округляем до следующего целого числа и получаем $L = 170$.

Где-то в далекой и холодной Антарктиде, на маленьком
 ледяном острове посреди черных₁ вод₂ океана₃ жил₄ пингвин₅.
В₆ этот₇ золотой₈ солнечный₉ день₁₀ в₁₁ своем₁₂ черном₁₃ меховом₁₄ фраке₁₅.
Пингвин₁₆ стоял₁₇ и₁₈ смотрел₁₉ далеко₂₀ вперед₂₁. Туда₂₂ где₂₃
встречаются₂₄ лазурное₂₅ небо₂₆ и₂₇ синяя₂₈ вода₂₉. Смотрел₃₀
и₃₁ думал₃₂ о₃₃ том₃₄ дне₃₅ когда₃₆ он₃₇ увидит₃₈ свою₃₉ пингвиниху₄₀
и₄₁ их₄₂ любимы₄₃ маленький₄₄ серый₄₅ комочек₄₆ с₄₇ черными₄₈
глазами₄₉. Но₅₀ вдруг₅₁ он₅₂ увидел₅₃ как₅₄ вводе₅₅ промелькнуло₅₆
что-то₅₇ красное₅₈. Это₅₉ была₆₀ стая₆₁ рыб₆₂. Как₆₃ только₆₄ он₆₅ это₆₆
увидел₆₇ так₆₈ сразу₆₉ бросился₇₀ в₇₁ воду₇₂ и₇₃ в₇₄ голову₇₅ покинули₇₆
все₇₇ мысли₇₈ кроме₇₉ еды₈₀. Еды₈₁ которая₈₂ так₈₃ необходима₈₄ ему₈₅
и₈₆ его₈₇ серому₈₈ комочку₈₉.

Рис. 7. Пример входного текста

На примере этого текста получается цветовая матрица в виде таблички (рис. 8).

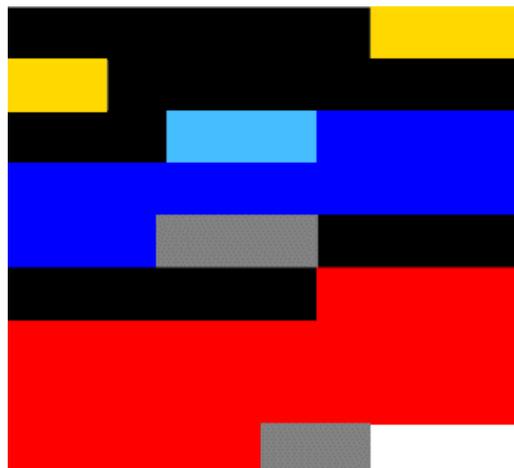


Рис. 8. Цветовая матрица входного текста

Алгоритм работы программы для создания цветовой матрицы следующий. Предполагается, что текст уже проанализирован и слова-цвета в нем найдены. Первый цвет: «*черных* вод океана...» – рисуем первый черный квадратик. Далее, от слова *черных* до слова *золотой* семь слов. Значит, программа рисует семь черных квадратиков. Восьмой квадратик уже будет золотого цвета. От слова *золотой* до слова *черном* у нас еще пять слов. Значит, система рисует пять золотых клеточек. Далее идут 11 черных, 3 лазурных, 17 синих, 3 серых клеточки и так далее до конца текста. Аналогичным образом была получена палитра для текста романа «Преступление и наказание» (см. рис. 6).

Помимо цветовой палитры программа может строить круговые диаграммы. Нажимаем Diagrams -> Get find colors diagram и на выходе получаем диаграмму цветов 1 (рис. 9). На данной диаграмме все цвета разделены по секторам. При наведении на сектор курсора мышки высказывает окошко с подсказкой, которое отображает количественную величину цвета. С помощью этой диаграммы можно наглядно увидеть, сколько и каких цветов было найдено в анализируемом тексте, и сравнить количество не только математически, но и визуально.

Нажимаем Diagrams -> Get images colors diagram и на выходе получаем диаграмму цветов 2 (рис. 10). Диаграмма 2 – это аналог цветовой матрицы, но в форме круговой диаграммы. Она показывает, в каком порядке и в каком объеме были найдены цвета в тексте.

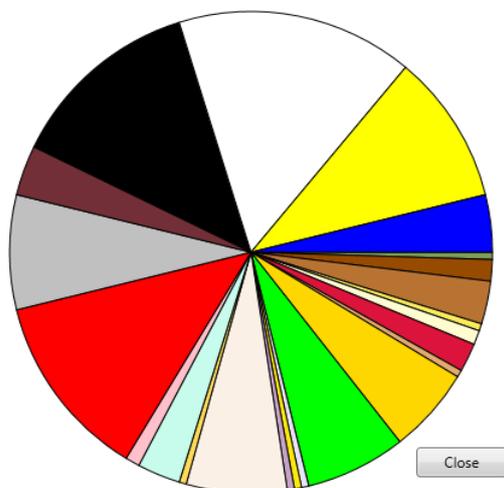


Рис. 9. Диаграмма цветов 1

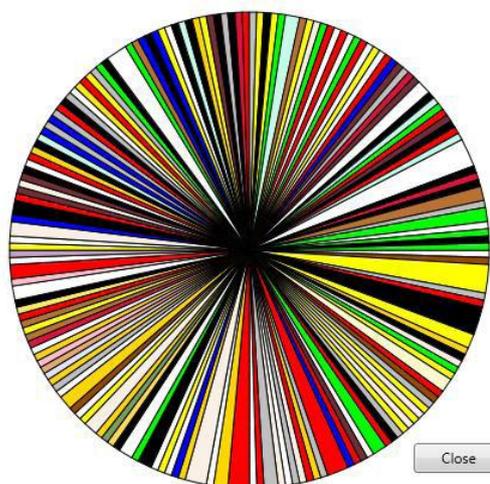


Рис. 10. Диаграмма цветов 2

При нажатии File -> Move to Excel статистика по 15 самым популярным цветам переносится в файл Excel. Это специальный файл, в котором при построении таблицы графики строятся цветами, которые были выявлены ранее. Помимо этого переносится общее значение цветов в тексте. Таким образом, можно найти разницу между всеми цветами и 15 самыми популярными. Можно проводить сравнительный анализ данных сразу нескольких книг. Это могут быть книги одного жанра, одного автора и многих других направлений. В итоге данная программа не только анализирует текст и рисует его цветовую матрицу и палитру, но и дает возможность работать со статистическими данными с помощью средств Excel (рис. 11).

```

1 | Sub SetChartColorsFromDataCells()
2 |
3 |   If TypeName(Selection) <> "ChartArea" Then
4 |     MsgBox "Сначала выделите диаграмму!"
5 |     Exit Sub
6 |   End If
7 |   Set c = ActiveChart
8 |   For j = 1 To c.SeriesCollection.Count
9 |     f = c.SeriesCollection(j).Formula
10 |    m = Split(f, ",")
11 |    Set r = Range(m(2))
12 |
13 |     For i = 1 To r.Cells.Count
14 |       c.SeriesCollection(j).Points(i).Format.Fill.ForeColor.RGB = _
15 |         r.Cells(i).Interior.Color
16 |     Next i
17 |   Next j
18 | End Sub

```

Рис. 11. Макрос для закрашивания графиков и диаграмм

После закрытия Visual Basic можно вернуться в Excel (рис. 12).

Использовать созданный макрос очень просто. Запуск макроса происходит с помощью кнопки Макросы на вкладке Разработчик (Developer – Macros) или с помощью сочетания клавиш Alt+F8. В том же окне можно в случае частого использования назначить макросу сочетание клавиш с помощью кнопки Параметры (Options).

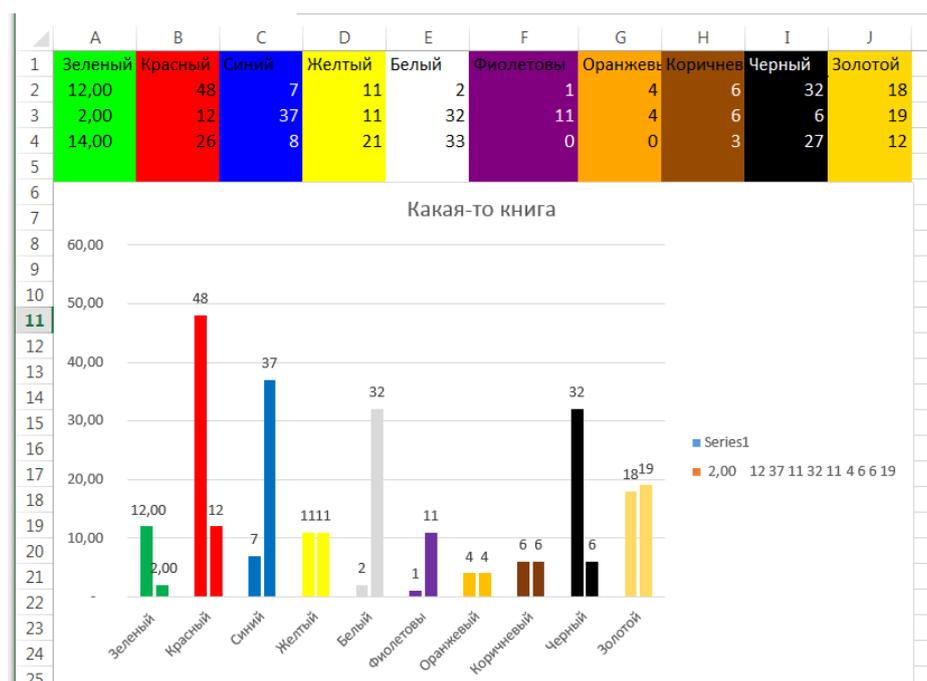


Рис. 12. Пример работы с созданным Excel-файлом

Разработанная система предназначена для распознавания образов на основе текстового содержания художественной литературы.

Заключение

Программа имеет две основные области применения: научно-исследовательскую и практическую. С точки зрения науки программа является средством для изучения и анализа влияния цвета на человека, его характер и психофизику путем разложения текстов на цветовые образы и сравнительного анализа числовых данных этих цветовых образов.

С практической точки зрения можно с помощью функции «цветовая матрица» создавать небольшие цветовые палитры, которые можно будет размещать на корешках книг как в обычных, так и интернет-магазинах. Благодаря «цветовой матрице» читатель сможет выбрать наиболее подходящую для него книгу на основании того, как выглядит рисунок, какие цвета и в каком объеме были использованы. Существует непосредственная зависимость между восприятием цветов книги и ее содержанием, сюжетом, историей, характером.

Помимо этого технологии данной программы можно применять в области психологии для изучения влияния книг и содержащихся в них цветов на подсознание человека и его внутренний мир.

Так как в настоящее время не существует непосредственных аналогов данного приложения, цветовой анализ может найти широкое применение в различных сферах человеческой жизнедеятельности.

Список литературы

1. Пескова, О.В. Алгоритмы классификации полнотекстовых документов / О.В. Пескова. – М. : МИЭМ, 2011. – 272 с.
2. Survey of Text Mining I: Clustering, Classification, and Retrieval / Ed. by M.W. Berry. – Springer, 2003. – 261 p.
3. Aggarwal, C.C. Mining Text Data / C.C. Aggarwal, C. Zhai. – Springer, 2012. – 527 p.
4. Пазельская, А. Метод определения эмоций в текстах на русском языке / А. Пазельская, А. Соловьев // The Intern. conf. on computational linguistics and intellectual technologies «Dialogue 2011». – М., 2011. – С. 510–522.

5. Жеребило, Т.В. Фоносемантика в словаре лингвистических терминов / Т.В. Жеребило. – 5-е изд. – Назрань : Пилигрим, 2010. – 486 с.

Поступила 21.09.2016

*Белорусский национальный
технический университет,
Минск, пр. Независимости, 65
e-mail: galarom@tut.by*

G.E. Romaniuk, M.S. Khodar

**THE ANALYSIS OF LEXICOGRAPHICAL COLORING
THE LITERARY TEXTS**

A text analysis program for finding and recognition of color characters used in the text is developed. The program fixes the data, builds a color matrix and a color diagram. The program exports the data into Excel for further study and analysis of statistics of color characters to determine their dependence on text type, genre and author.

УДК 004.91

Л.В. Серебряная, В.В. Потараев

МЕТОДЫ КЛАССИФИКАЦИИ ТЕКСТОВОЙ ИНФОРМАЦИИ НА ОСНОВЕ ИСКУССТВЕННОЙ НЕЙРОННОЙ И СЕМАНТИЧЕСКОЙ СЕТЕЙ

Рассматривается применение искусственной нейронной сети в виде перцептрона, сети Хопфилда и семантической сети для классификации текстовой информации. Изучаются процедуры обучения сетей, реализуются алгоритм обратного распространения ошибки в перцептроне и алгоритм сходимости сети Хопфилда. Предлагается программное средство автоматической классификации текстов на основе разработанных моделей и алгоритмов. Оцениваются результаты работы программного средства.

Введение

В последние десятилетия резко возросла актуальность работы с документами, представленными в цифровом виде. Вследствие этого появились новые модели, методы и инструменты для работы с ними.

Хранение больших объемов информации практически оправдано только при условии, если ее поиск и обработка осуществляются быстро и выдается она в доступной для понимания форме. Один из видов обработки текстовой информации – это ее классификация (или рубрикация), которая находит ряд применений: фильтрация спама, разделение электронных сообщений по категориям, подбор контекстной рекламы, классификация научных статей и др.

Классификации информации предшествует этап ее анализа. Большинство существующих подходов к анализу текстов можно разбить на два типа. К первому относятся простые, быстрые, но не очень точные механизмы анализа. Чаще всего эти подходы используют формальные статистические методы, основанные на частоте появления в тексте слов различных тематик. Второй тип формируют достаточно сложные, дающие хороший результат, но сравнительно медленные подходы, основанные на лингвистических методах. Эффективным же можно считать такой подход, который сочетал бы в себе «простоту» статистических алгоритмов с достаточно высоким качеством обработки лингвистических методов.

Одним из видов лингвистического анализа является семантический анализ. С его помощью исследуются смысл слов и предложений, их связь между собой и с окружающей действительностью. Методы реализации семантического анализа и классификации текстовой информации связаны с использованием искусственных нейронных сетей (ИНС) и семантических сетей, которым посвящена данная работа.

1. Модели обработки текстовой информации на основе семантического анализа

Накопленный теоретический и практический опыт работы с ИНС позволяет выделить множество вариантов их построения, рассмотренных ниже.

Если ИНС состоит из нейронов одного типа, она называется однородной, если же в ней комбинируются слои нейронов разных типов, – гибридной.

По количеству слоев нейронов сети делятся на однослойные, в которых отсутствуют непосредственные связи выходов одних нейронов со входами других, и многослойные, где имеются указанные связи.

Сеть является однонаправленной, если в ней отсутствует передача сигналов с последующих слоев на предыдущие. Нейронная сеть с обратными связями называется рекуррентной.

По типу обучения ИНС делятся на обучающиеся с учителем (контролируемое обучение) и самообучающиеся. Для сетей первого типа задается обучающая выборка, для каждого образа которой заранее известна принадлежность к одному из заданных классов. В сетях второго типа,

популярным представителем которых считается сеть Кохонена, обучающая выборка отсутствует, а определение количества классов и отнесение к ним образов выполняются в ходе реализации алгоритма распознавания.

Многослойные сети могут привести к увеличению вычислительной мощности по сравнению с однослойной сетью лишь в том случае, если активационная функция между слоями будет нелинейной. Иначе любая многослойная линейная сеть может быть заменена эквивалентной однослойной сетью, которая весьма ограничена по своим вычислительным возможностям.

Анализ различных типов ИНС показал, что для классификации текстовой информации наиболее подходят два из них: многослойный персептрон и сеть Хопфилда. Рассмотрим их более подробно [1].

Одним из самых важных условий успешной работы ИНС является алгоритм обучения сети. Многослойный персептрон (рис. 1) обучается с помощью алгоритма обратного распространения ошибки. Целью обучения сети этим алгоритмом является такая корректировка ее коэффициентов, при которой инициализация множества входов приводит к требуемому множеству выходов. Обычно множества входов и выходов называются векторами X и Y соответственно. При обучении предполагается, что для каждого входного вектора существует парный ему целевой вектор, задающий требуемый выход. Вместе они образуют обучающую пару. Сеть обучается на многих парах.

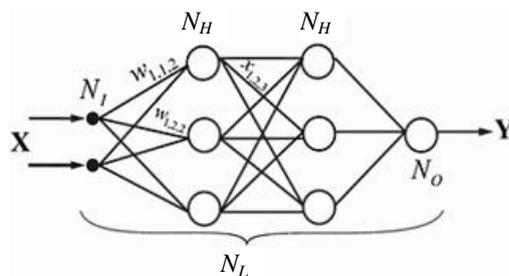


Рис. 1. Сеть в виде многослойного персептрона

Алгоритм обратного распространения ошибки предполагает два прохода по всем слоям сети: прямой и обратный. При прямом проходе входной вектор подается на входной слой нейронной сети, после чего распространяется по сети от слоя к слою. В результате генерируется набор выходных сигналов, который и является фактической реакцией сети на данный входной образ. Во время прямого прохода все весовые коэффициенты сети фиксированы. Во время обратного прохода все коэффициенты настраиваются в соответствии с правилом коррекции ошибок, а именно: фактический выход сети вычитается из желаемого, в результате чего формируется сигнал ошибки. Впоследствии этот сигнал распространяется по сети в направлении, обратном направлению связей весовых коэффициентов. Они настраиваются с целью максимального приближения выходного сигнала сети к желаемому [2].

В качестве активационной функции в многослойных персептронах, как правило, используется сигмоидальная функция, в частности бинарный сигмоид:

$$f(x) = \frac{1}{1+e^{-x}}. \quad (1)$$

Выражение

$$f'(x) = f(x) * [1 - f(x)] \quad (2)$$

является производной функцией бинарного сигмоида.

Алгоритм обратного распространения ошибки состоит из следующих шагов:

1. Инициализировать весовые коэффициенты маленькими случайными значениями.
2. Выбрать очередную обучающую пару из обучающего множества, подать входной вектор на вход сети.
3. Вычислить выход сети.

4. Вычислить разность между выходом сети и требуемым выходом (целевым вектором обучающей пары).

5. Подкорректировать коэффициенты сети для минимизации ошибки.

6. Повторять шаги 2–5 для каждого вектора обучающего множества до тех пор, пока ошибка на всем множестве не достигнет приемлемого значения.

Операции, выполняемые на шагах 2 и 3, сходны с теми, которые выполняются при функционировании уже обученной сети, т. е. подается входной вектор и вычисляется получающийся выход. Вычисления выполняются послойно. Шаги 2 и 3 образуют «проход вперед», так как сигнал распространяется по сети от входа к выходу. Шаги 4 и 5 составляют «обратный проход», когда вычисляемый сигнал ошибки распространяется обратно по сети и используется для корректировки весовых коэффициентов.

Рассмотрим подробнее процедуру корректировки коэффициентов ИНС, при которой следует выделить два возможных случая:

1. Изменение коэффициентов выходного слоя.

Введем величину δ , которая равна разности между требуемым T и реальным OUT выходами, умноженной на производную функции активации (бинарный сигмоид):

$$\delta_q = OUT_q(1 - OUT_q)(T_q - OUT_q). \quad (3)$$

Тогда коэффициенты выходного слоя после коррекции определяются по формуле

$$\omega_{p-q}(i+1) = \omega_{p-q}(i) + \eta \delta_q OUT_p, \quad (4)$$

где i – номер текущей итерации обучения;

ω_{p-q} – величина весового коэффициента, соединяющего нейрон p с нейроном q ;

η – коэффициент скорости обучения, позволяющий управлять средней величиной изменения коэффициентов.

2. Корректировка весовых коэффициентов скрытого слоя.

Введем величину δ :

$$\delta_q = OUT_q(1 - OUT_q) \sum_{k=1}^M \delta_k \omega_{q-k}. \quad (5)$$

Коэффициенты скрытых слоев после коррекции также будут определяться с помощью выражения (4).

Достоинством алгоритма обратного распространения ошибки является то, что он реализует вычислительно эффективный метод обучения многослойного персептрона. При этом его все-таки нельзя считать универсальным решением. Больше всего трудностей возникает с неопределенно долгим процессом обучения, в ходе которого величина шага коррекции коэффициентов может конфликтовать с качеством и временем обучения. В роли входных значений обучающей выборки выступают наборы ключевых слов текстов, а в роли выходных значений – номера классов, к которым относятся эти же тексты.

Сеть Хопфилда занимает особое место в ряду ИНС. В ней впервые удалось установить связь между нелинейными динамическими системами и нейронными сетями. Образы памяти сети соответствуют устойчивым предельным точкам (аттракторам) динамической системы. При этом появилась возможность теоретически оценить объем памяти сети Хопфилда, определить область ее параметров, в которой достигается наилучшее функционирование [3].

В общем случае модель Хопфилда может быть представлена сетью, содержащей произвольные обратные связи. По этим связям переданное возбуждение возвращается к данному нейрону, и он повторно выполняет свою функцию (рис. 2).

Обратные связи могут вызывать неустойчивости в поведении ИНС, что проявляется в блуждающей смене состояний нейронов, не приводящей к стационарным состояниям. Однако было показано, что сети Хопфилда устойчивы, и их можно определить как динамическую систему с обратной связью, у которой выход одной операции служит входом следующей операции сети. Каждая операция сети называется итерацией. Устойчивость сети подразумевает, что она

может сходиться к одной из зафиксированных (неподвижных) точек, которая зависит от исходной точки, выбранной для начальной итерации. Множество неподвижных точек сети Хопфилда – это ее память. В этом случае сеть может действовать как ассоциативная память.

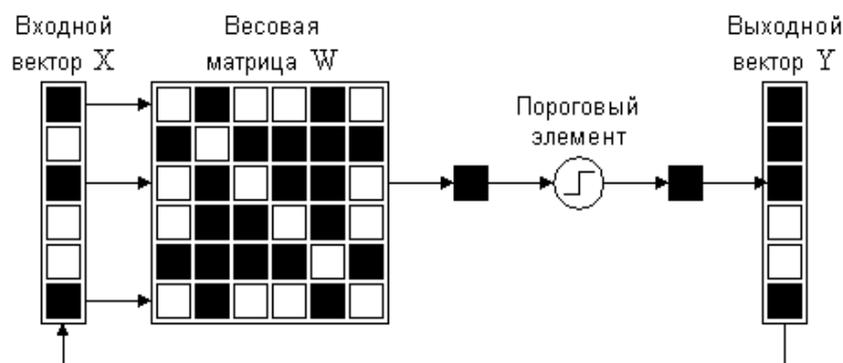


Рис. 2. Нейронная сеть Хопфилда

Рассмотрим алгоритм сходимости сети Хопфилда:

1. Вычислить компоненты выходного вектора $Y_j, j = 1, 2, \dots, n$, по формуле

$$Y_j = T(\sum_{i=1}^n w_{ij} x_i), \quad (6)$$

где $T(x) = \text{sign}(x)$, т. е.

$$T(x) = \begin{cases} -1, & \text{если } x < 0; \\ 1, & \text{если } x > 0. \end{cases} \quad (7)$$

2. Выполнить асинхронную коррекцию.

- 2.1. По формуле (6) найти вектор Y_j .

- 2.2. Заменить $X = (x_1, x_2, x_3, \dots, x_n)$ на $Y = (y_1, x_2, x_3, \dots, x_n)$ и подать Y на вход X .

- 2.3. Повторить процесс, чтобы найти y_2, y_3 и т. д.

- 2.4. Повторять шаги 2.2 и 2.3 до тех пор, пока вектор не перестанет изменяться. Каждый шаг уменьшает величину энергии связей, поэтому обеспечивается сходимость к неподвижной точке (аттрактору).

Асинхронная коррекция и нули на диагонали матрицы весовых коэффициентов W гарантируют, что энергетическая функция будет уменьшаться с каждой итерацией. Весовая матрица отличает поведение одной сети Хопфилда от поведения другой.

К недостаткам сети Хопфилда можно отнести ее сравнительно небольшой объем памяти, вследствие чего попытка записи большего числа образов приводит к тому, что нейронная сеть перестает их распознавать. Кроме того, достижение устойчивого состояния не гарантирует правильный ответ сети. Это происходит из-за того, что ИНС может сойтись к так называемым ложным аттракторам.

Для обучения сети Хопфилда составляется обучающая выборка. Это наборы ключевых слов знакомых текстов. Наличие слова в тексте означает единичный сигнал на входе нейронной сети. При обучении можно рассчитать коэффициенты матрицы связей на основании слов обучающих текстов. При классификации текстов входной сигнал после выполнения некоторого количества итераций будет приводить нейронную сеть в стационарное состояние, соответствующее образу, на котором проводилось обучение. Каждый из обучающих образов соответствует одному из классов, к которым нужно относить незнакомые тексты.

Еще одним способом семантической обработки текстов является подход на основе семантических сетей. Семантические сети позволяют выделять смысл текста в виде понятий и связей между ними, образующих граф. Понятия семантической сети записываются в вершинах графа, а отношения между понятиями – это дуги графа. Количество типов отношений в семан-

тической сети определяется ее разработчиком исходя из конкретных целей. Часто используются иерархические семантические сети, в которых отношения образуют древовидную структуру. Отношения в сетях могут быть разных типов: функциональными, количественными, пространственными, временными, логическими и др. [4].

К достоинствам семантических сетей можно отнести:

- универсальность, достигаемую за счет выбора соответствующего набора отношений;
- наглядность системы знаний, представленную графически;
- близость структуры сети, представляющей систему знаний, к семантической структуре фраз на естественном языке;
- соответствие современным представлениям об организации долговременной памяти человека.

Недостатки семантических сетей:

- сетевая модель не всегда дает ясное представление о структуре предметной области, поэтому формирование и модификация такой модели могут быть затруднительными;
- сетевые модели представляют собой структуры, для обработки которых необходим специальный аппарат формального вывода;
- проблема поиска решения в семантической сети сводится к задаче поиска ее фрагмента, отражающего поставленный запрос. Это может обуславливать сложность поиска решения в семантических сетях;
- представление, использование и модификация знаний при описании систем реального уровня сложности оказываются трудоемкими процедурами, особенно при наличии множественных отношений между их понятиями.

При решении задачи классификации текстовой информации с помощью семантических сетей можно в качестве узлов сети принять некоторые концепты, о которых идет речь в тексте, а в качестве дуг – связи между концептами. В таком случае обучение классификатора будет представлять собой создание сетей по некоторым обучающим текстам – одна сеть на одну рубрику. При классификации текстов для каждого из них нужно построить собственную семантическую сеть, а затем на основании сходства сети текста и сетей рубрик относить текст к одной из рубрик.

2. Алгоритмы классификации текстов с помощью искусственных нейронных и семантических сетей

Рассмотрим алгоритмы классификации текстов на основе ИНС двух типов: персептрона и сети Хопфилда. В обоих случаях выполняется контролируемое обучение сетей, для чего используются наборы ключевых слов обучающих текстов, построенные по законам Зипфа.

Для выделения понятий текста, представляющих слова и словосочетания, может быть применен статистический алгоритм, основанный на анализе частоты встречаемости слов, цепочек слов и их вхождения друг в друга. Во всех созданных человеком текстах можно выделить статистические закономерности, которые никому не удастся обойти. Независимо от текста и языка его написания внутренняя структура текста остается неизменной.

Если измерить количество вхождений каждого слова в текст и взять только одно значение из каждой группы, расположить частоты по мере их убывания и пронумеровать (порядковый номер частоты называется рангом частоты), то наиболее часто встречающиеся слова будут иметь ранг 1, следующие за ними – 2 и т. д. Кривая Зипфа, имеющая вид равносторонней гиперболы, отражает зависимость между частотой вхождения слов в текст и рангом слов. Исследования показали, что наиболее значимые слова лежат в средней части кривой и имеют ранг от 4 до 17. Это объясняется тем, что слова, которые встречаются очень часто, в основном оказываются предлогами и местоимениями. Редко встречающиеся слова в большинстве случаев тоже не имеют решающего смыслового значения [5].

Количество входных нейронов сети равно суммарному количеству выбранных ключевых слов во всех текстах:

$$N = M \times K, \quad (8)$$

где N – количество входов нейронной сети; M – количество слов, отобранных из каждого текста; K – количество обучающих текстов.

Выходам нейронной сети можно сопоставить классы, к которым относятся обучающие тексты.

Рассмотрим процедуру классификации текстов при помощи алгоритма обратного распространения ошибки.

В качестве функции активации нейронов используется бинарный сигмоид (1) и его производная (2). Начальным значениям весовых коэффициентов связей между нейронами присваиваются случайные числа от 0 до 1. Скорость обучения подбирается таким образом, чтобы оно завершалось в течение нескольких итераций обратного распространения ошибки. Еще одним параметром является максимально допустимое количество итераций обучения.

По завершении обучения выполняется проверка работы сети на обучающей выборке. Каждое из M ключевых слов каждого из K текстов подается на входы сети, и она должна корректно определить номер класса. Только в этом случае обучение считается успешным.

Для проверки классификации используется набор тестовых текстов, для каждого из которых сеть должна определить выход (класс принадлежности). Если такие выходы не найдены или их найдено несколько для одного текста, то решение о классификации не принимается. В то время как процесс обучения является итерационным, классификация представляет собой расчет функции активации для каждого из узлов нейронной сети.

В случае выполнения алгоритма классификации текстов на основе нейронной сети Хопфилда обычно используется функция активации, представленная выражением (7). Начальные весовые коэффициенты связей нейронов задаются случайными небольшими значениями, а процесс обучения заключается в расчете матрицы связей W .

После расчета матрицы связей нейронная сеть готова к выполнению классификации. Она представляет собой итерационный процесс, при котором сеть последовательно меняет свои состояния. Это происходит до тех пор, пока ИНС не попадет в некоторое стационарное состояние, соответствующее одному из обучающих сигналов. Поскольку условие прекращения итераций может никогда не выполниться, классификацию обычно ограничивают максимальным числом итераций. Если по окончании итераций состояние сети не соответствует ни одному из обучающих текстов, то классифицируемый текст относят к классу, наиболее близкому к одному из обучающих сигналов.

Для классификации текстов на основе семантической сети представим текст как множество предложений, каждое из которых состоит из подлежащего, сказуемого и дополнений. Узлы сети – это некоторые понятия, выражаемые подлежащими и дополнениями, а связи между узлами задаются сказуемыми. С целью упрощения решаемой задачи остальные члены предложений исключены из рассмотрения. Для определения части речи слов, а также их начальных форм используются словари начальных форм существительных и глаголов русского языка. Особенностью семантической сети является определение схожести текстов по их смыслу, даже если в них используются разные слова. Для обеспечения такой возможности узлом сети принято считать не только само слово, но и все его синонимы. Аналогично связью можно считать не глагол, а множество синонимичных глаголов, означающих примерно одинаковые отношения.

Рассмотрим алгоритм классификации текстов, использующий сеть, построенную предложенным способом. На этапе обучения происходит добавление каждого предложения текста к создаваемой семантической сети. Если подлежащее (либо дополнение) уже есть в сети, то к нему добавляется дуга, соответствующая сказуемому. Если в предложении нет сказуемого, то оно со всеми своими синонимами добавляется в сеть как не связанный ни с чем узел. На рис. 3 показан пример семантической сети, построенной для двух предложений.

Назовем участком сети любой узел либо дугу и пару связанных с ней узлов. На этапе классификации при поиске участка сети, наиболее близкого к предложению классифицируемого текста, близость предложения и участка сети можно вычислять по-разному. Например, если совпадают подлежащие в двух дугах, то близость равна X , а если еще совпадают сказуемые, то близость равна $X + Y$. Будем считать величины X и Y весами близости подлежащего и сказуемого соответственно.

Необходимо отметить, что количество итераций данного алгоритма зависит только от числа предложений в обрабатываемых текстах. Близость текстов можно оценивать на основе оценки близости дуг соответствующих им сетей.

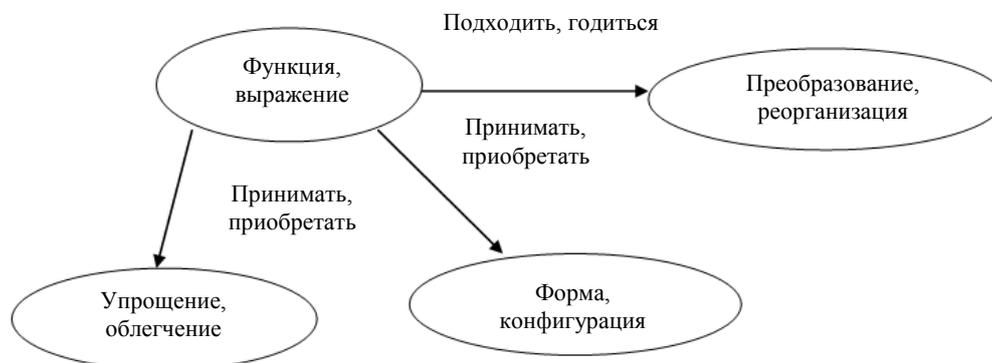


Рис. 3. Фрагмент семантической сети

Для работы алгоритма нужны словарь начальных форм существительных, словарь глаголов и словарь синонимов русского языка. На начальном этапе для ускорения алгоритма можно заменить каждый текст подмножеством его предложений (например, 100 первых предложений текста или 100 самых длинных предложений). Данную сеть можно дополнить и другими правилами, которые позволяют генерировать новые высказывания, увеличивая порождающую способность семантической сети. Тогда для повышения качества работы сети вводится вероятность каждой связи.

3. Сравнение методов классификации текстов и анализ полученных результатов

На основе ранее рассмотренных моделей и алгоритмов было разработано программное средство, предназначенное для автоматической классификации текстовых документов.

На персептроне и сети Хопфилда проводилась процедура контролируемого обучения. Для этого на входы сетей подавались векторы ключевых слов текстов, составляющих обучающую выборку. Количество классов, размер обучающей выборки, принадлежность текстов к рубрикам задает пользователь, при этом все данные могут изменяться. Результат обучения – определение номера класса для каждого обучающего образа. В качестве начальных значений весовых коэффициентов связей между нейронами использовались случайные числа от 0 до 1. Скорость обучения подбиралась таким образом, чтобы для небольшой обучающей выборки процесс обучения завершался достаточно быстро. Дополнительно можно задать максимальное количество итераций обучения.

По завершении процедуры обучения можно оценить характеристики классификации каждой из ИНС. Сначала задавались три класса, десять обучающих текстов, каждый из которых был представлен вектором из десяти ключевых слов. Размер текстов – 100–200 Кб. Качество классификации текстов оказалось у сетей приблизительно одинаковым, но сеть Хопфилда работала быстрее. С увеличением обучающей выборки и вектора ключевых слов увеличивалось время обучения, а качество классификации становилось выше.

Перейдем к реализации семантической сети. В простейшем случае в разработанной сети каждому предложению текста ставится в соответствие некоторая дуга. Близость текстов фактически означает совпадение дуги либо узла сети одного текста с дугой либо узлом другого текста. В качестве параметров сети выбирались следующие: минимальная близость для принятия решения о классификации; веса близости подлежащих, сказуемых, дополнений. Для заданных трех классов семантическая сеть выдала более 80 % корректных ответов, что явилось лучшим результатом, чем результаты обеих ИНС. Время работы семантической сети оказалось больше, чем у нейронных сетей.

Рассмотрим пример разделения текстов на пять классов с помощью двух ИНС и семантической сети (таблица). С увеличением числа классов качество классификации перцептроном немного ухудшилось, сетью Хопфилда – стало существенно хуже, а семантическая сеть работала практически без изменений. Это можно объяснить следующим образом: перцептрон не хватило итераций для того, чтобы закончить процесс обучения, а сеть Хопфилда начала сходиться к ложной стационарной точке. Кроме того, с увеличением обучающей выборки и длины векторов ключевых слов ИНС для обучения требуется большее количество итераций. В результате они теряют преимущество по времени перед семантической сетью.

Результаты классификации текстов на пять рубрик

Количество текстов	Рубрика текста	Количество текстов, использованных для обучения	Количество верных ответов		
			Классификация перцептроном	Классификация сетью Хопфилда	Классификация семантической сетью
7	Теория вероятностей	2	6	6	5
12	Философия	3	9	7	6
15	Информатика	3	12	8	10
20	Геометрия	4	15	10	17
22	Физика	3	15	10	20

Созданное программное средство позволяет подбирать оптимальные параметры сетей в зависимости от требований пользователя. Так, например, при небольшом количестве классов и нежестких требованиях к качеству классификации можно выбирать сеть Хопфилда. Алгоритм на основе нейронной сети с обратным распространением ошибки показывает достаточно стабильные результаты при различных количествах классов, допуская много ошибок. Для достижения баланса между скоростью и качеством классификации можно использовать две сети: нейронную сеть с обратным распространением ошибок и предложенную семантическую сеть в случае, если нейронная сеть не сумела определить класс.

Заключение

В работе изучены модели и алгоритмы классификации текстовой информации. Предложено применение методов семантического анализа текстовой информации для решения задачи классификации. Семантический анализ связан с выделением информационно-логической основы текста, что и было выполнено в работе.

Предложены алгоритмы классификации текстовой информации на основе искусственных нейронных и семантической сетей. Для всех моделей применялся метод контролируемого обучения, обеспечивающий более точное решение поставленной задачи.

Создано программное средство, реализующее алгоритм обратного распространения ошибки для перцептрона, алгоритм сходимости сети Хопфилда и алгоритм классификации семантической сети. Работу программного средства можно настраивать с помощью различных параметров в зависимости от решаемой прикладной задачи. Результаты работы программы представляются в удобном аналитическом и графическом виде. Сравнительный анализ полученных результатов показал, что семантическая сеть дает более точные результаты, хотя и отстает от ИНС по скорости работы. При увеличении количества классов разница в скорости работы сетей уменьшается. Для улучшения результатов по качеству и времени работы предлагается комбинировать различные модели и алгоритмы сетей.

Список литературы

1. Искусственная нейронная сеть [Электронный ресурс]. – Режим доступа : https://ru.wikipedia.org/wiki/Искусственная_нейронная_сеть. – Дата доступа : 06.06.2016.

2. Алгоритм обратного распространения ошибки [Электронный ресурс]. – Режим доступа : <http://www.aiportal.ru/articles/neural-networks/back-propagation.html>. – Дата доступа : 06.06.2016.
3. Нейронная сеть Хопфилда и ее применение [Электронный ресурс]. – Режим доступа : <http://iasa.org.ua/lections/tpr/neuro/hopfield.htm>. – Дата доступа : 06.06.2016.
4. Семантические сети или сетевые модели знаний [Электронный ресурс]. – Режим доступа : <http://www.aiportal.ru/articles/knowledge-models/semantic-network.html>. – Дата доступа : 06.06.2016.
5. Серебряная, Л.В. Информационное обеспечение финансовых структур / Л.В. Серебряная // Методическое пособие к лабораторным работам для студентов специальности «Программное обеспечение информационных технологий» всех форм обучения. – Минск : БГУИР, 2011. – 43 с.

Поступила 13.07.2016

*Белорусский государственный университет
информатики и радиоэлектроники,
Минск, ул. П. Бровки, 6
e-mail: l_silver@mail.ru,
vic229@rambler.ru*

L.V. Serebryanaya, V.V. Potaraev

METHODS OF TEXT INFORMATION CLASSIFICATION ON THE BASIS OF ARTIFICIAL NEURAL AND SEMANTIC NETWORKS

The article covers the use of perceptron, Hopfield artificial neural network and semantic network for classification of text information. Network training algorithms are studied. An algorithm of inverse mistake spreading for perceptron network and convergence algorithm for Hopfield network are implemented. On the basis of the offered models and algorithms automatic text classification software is developed and its operation results are evaluated.

УДК 65.011.56; 007.51

И.В. Шатохин

ИНФОРМАЦИОННАЯ СИСТЕМА УПРАВЛЕНИЯ ПРЕДПРИЯТИЕМ КАК ЦЕПОЧКОЙ ПОСТАВОК

Предлагается подход к построению автоматизированной системы управления логистикой предприятия, основанный на использовании систем синхронного планирования цепочек поставок. В процессе планирования рассчитываются графики производства и отгрузки товаров в сбытовые точки, сбалансированные с доступностью материалов и производственными мощностями. Производится прогноз дефицита в обеспечении производства покупными материалами и комплектующими изделиями с целью его оперативного предупреждения. Обеспечивается контроль за ходом выполнения заказов на закупку, а также рассматривается структура системы автоматизации управления цепочками поставок, ее цели и задачи.

Введение

Рассмотрим предприятие, где сбытовая сеть представлена несколькими территориально распределенными сбытовыми точками. Сбытовые точки обеспечиваются товарами за счет собственного производства предприятия, которое территориально удалено от них. Перемещения изделий между сбытовыми точками и возвраты изделий из сбытовых точек на производство отсутствуют. Цикл производства изделий короткий (несколько дней). Будем считать, что с точки зрения планирования производства изделия производятся мгновенно с одновременным потреблением заданного физического объема некоторого перечня покупных комплектующих изделий (ПКИ) и материалов. Выпуск изделий ограничивается производственными мощностями. Динамика фонда времени работы оборудования определяется не только профилактическими мероприятиями (ремонт и профилактическим обслуживанием технологического оборудования), но и вводом новых мощностей в эксплуатацию. Динамика фонда времени основных работников определяется как графиком отпусков, текучестью, так и мероприятиями по регулированию численности трудовых ресурсов.

Перечень производимых изделий является динамическим. Часть изделий с заданного периода планирования перестает производиться, тогда принимается решение о прекращении их выпуска. Новые изделия запускаются в производство с определенного момента времени. В процессе производства вводятся новые (несерийные) изделия со своим нормативным потреблением комплектующих изделий и материалов. Выпущенные несерийные изделия не включаются в запас готовой продукции предприятия и не отгружаются в сбытовые точки. Для них задается не план продаж, а план производства.

1. Планирование цепочек поставок

Для вышеописанного предприятия производственное планирование можно рассматривать в контексте управления цепочками поставок с учетом ряда ограничений. Цепочка поставок есть взаимосвязанная система отношений между поставщиками материалов и услуг, которая охватывает весь цикл трансформации материальных ресурсов из исходного сырья и материалов в конечные продукты и услуги, а также доставку этих продуктов и услуг конечному потребителю [1, 2].

Модель информационных потоков в данном случае практически вписывается в структурную формулу взаимосвязей источников и потребителей информации: *сбытовая сеть – производство – поставщики комплектующих*, т. е. информационные потоки противоположны материальным (рис. 1). В соответствии с этим *сбытовая сеть* порождает планы продаж, которые должны быть реализованы за счет производства товаров. В свою очередь, производство должно быть своевременно обеспечено материалами и комплектующими изделиями, которые закупаются у поставщиков. *Сбытовая сеть* может порождать несколько вариантов планов продаж (оптимистичный, пессимистичный и т. п.), которые с течением времени подвергаются корректировке.

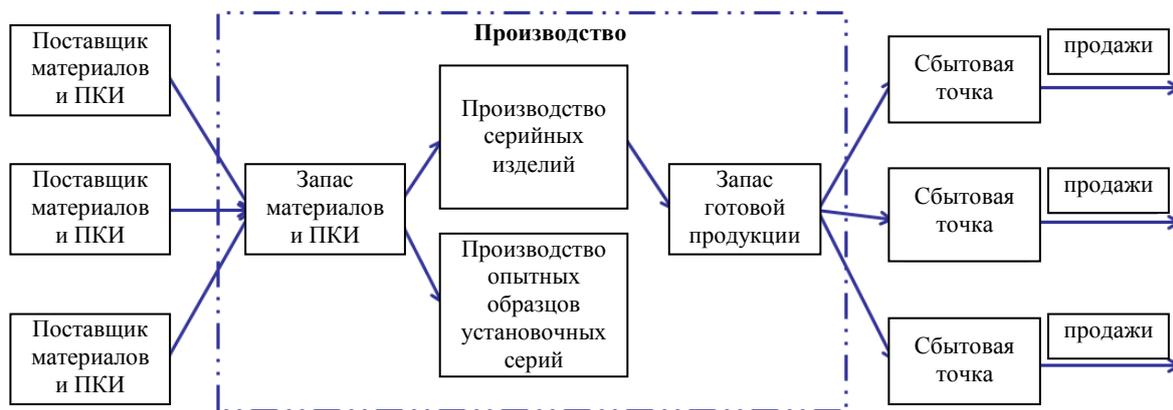


Рис. 1. Модель материальных потоков предприятия

Основные ограничения связаны с дефицитом производственных мощностей и трудовых ресурсов, а также с закупкой материалов и комплектующих изделий. Если две первые проблемы требуют для своего решения вливания серьезных финансовых средств, связанных с наймом рабочей силы и техническим перевооружением производства, то ограничение с закупкой комплектующих часто снимается созданием пула альтернативных поставщиков, каждый из которых может обеспечивать поставки по одному или нескольким альтернативным маршрутам. В пул включаются поставщики (процесс сертификации), отвечающие требуемым критериям по цене и качеству поставляемых ими материалов и комплектующих изделий. Обычно таких поставщиков мало (два-три, один из которых основной, а остальные резервные). Во многих случаях по ответственной номенклатуре существует вообще единственный поставщик-монополист.

Задача планирования сводится к подбору такого варианта плана производства изделий и закупок материалов, который позволит уложиться в заданные ограничения по мощности и сроки поставок материалов с максимально возможным обеспечением плана продаж на заданном горизонте планирования. При решении подобной задачи необходимо сопоставить и увязать большой объем данных о работе предприятия. Естественно, проделать такую работу в разумные сроки невозможно без автоматизированной системы планирования.

Наилучшим образом изложенным выше требованиям соответствуют системы синхронного планирования, в основе которых лежит APS-алгоритм (Advanced Planning and Scheduling Algorithm) [3]. В процессе планирования алгоритм синхронизирует загрузку доступных мощностей с планами движения материалов и комплектующих, что позволяет оценить реалистичность планов продаж с отражением возможных отклонений. Суть работы алгоритма заключается в том, что на основе анализа графика продаж товаров в сбытовых точках и обеспеченности его наличными запасами и товарами, находящимися в пути, он формирует заказы на отгрузку продукции с завода в сбытовые точки. Эти заказы в случае их выполнения должны ликвидировать плановый дефицит продаж. Объемы заказов на отгрузку могут рассчитываться как на основе среднесуточных объемов продаж товаров, так и в соответствии с заданной минимальной партией отгрузки. При необходимости заказы рассчитываются с учетом обеспечения нормативного переходящего запаса товаров с одного планового периода на другой (страхового запаса).

Далее осуществляется вовлечение запасов готовой продукции на складе завода в счет обеспечения сформированных заказов на отгрузку, а на оставшуюся часть формируются заказы на производство, исполнение которых планируется во времени с учетом доступности мощностей и трудовых ресурсов. Объемы заказов на производство вычисляются по нормативам на размеры партий запуска-выпуска продукции и с учетом нормативного коэффициента выхода годных.

На формирование очередности исполнения заказов могут влиять следующие параметры:

- приоритеты производства номенклатурных позиций готовой продукции;
- время опережения запуска заказов, которое определяет, насколько раньше можно начать производство продукции, чтобы максимально обеспечить потребности сбытовых точек в заданном плановом периоде.

Под полученное расписание исполнения заказов на основе норм рассчитываются потребности в материалах ПКИ и формируются заказы на их закупку таким образом, чтобы они поступили точно в срок с учетом норм длительности доставки и размещения заказов у поставщиков. Норма длительности размещения заказа определяется длительностью производственного цикла поставщика с момента получения им заявки заказчика до момента отгрузки продукции в адрес грузополучателя. При этом также возможно внесение поправок в нормативные сроки оформления и согласования заказа с поставщиком и нахождение материалов на входном контроле предприятия.

Расчет потребности в материалах и комплектующих изделиях проводится с учетом нормативных потерь покупных материалов и ПКИ в производстве. При формировании заказов на закупку учитывается:

- наличие запасов материалов и ПКИ в производстве;
- плановое поступление материально-технических ресурсов от поставщиков по уже размещенным заказам;
- ожидаемый процент брака комплектующих изделий на входном контроле.

Для гарантии своевременных поставок система подбирает поставщиков и маршруты доставки материалов исходя из критерия минимизации стоимости транспортно-заготовительных расходов и сроков доставки. Часто в таких условиях оперативная работа службы снабжения в процессе подбора поставщика из имеющегося пула сводится к анализу требуемой скорости и стоимости доставки (выбора маршрута), чтобы к моменту возникновения потребности необходимые материалы были на складе предприятия. Сначала делается расчет по основному поставщику. Если он не дает требуемого результата, то пользователь переключается на расчет по резервному поставщику из пула. В итоге формируется сеть взаимосвязанных заказов, где каждый предыдущий заказ порождает один или несколько последующих заказов. В процессе расчета контролируется дата запуска заказа, чтобы она «не ушла в прошлое» за дату начала планового периода. Если такое происходит, алгоритм автоматически смещает все взаимосвязанные заказы в будущее.

2. Структура и задачи системы автоматизации управления

Как правило, система автоматизации управления цепочками поставок включает в себя подсистемы ведения нормативно-справочной информации (НСИ), управления логистикой и складского учета и планирования производства изделий и закупок материалов.

Подсистема ведения НСИ должна обеспечивать возможность ведения данных:

- о нормативах станкоемкости и трудоемкости производимой продукции;
- сводных нормах расхода материалов и комплектующих изделий с учетом их потерь при производстве продукции;
- графиках работы технологического оборудования и основного производственного персонала;
- допустимых заменах материалов и комплектующих изделий в производстве (ведомость замен).

Подсистема управления логистикой и складского учета является источником информации:

- о наличии запасов товаров в сбытовых точках и на складах готовой продукции;
- товарах, находящихся в пути, к сбытовым точкам;
- сбытовых точках (нормативах длительности доставки в них товаров, минимальных партиях отгрузки, объемах страховых запасов и т. п.);
- наличных запасах материалов и комплектующих изделий с учетом сроков их годности и гарантийного хранения (последнее реализуется, например, средствами партионного учета материально-производственных запасов);
- материалах и комплектующих изделиях, находящихся в пути;
- поставщиках (поставляемой номенклатуре; нормативных сроках размещения заказов; параметрах маршрутов поставки, таких как длительность доставки, нормативы транспортно-заготовительных затрат; ожидаемом проценте брака; размерах минимальных партий поставки и их кратности, ценах на материалы и комплектующие изделия и т. п.);

– размерах минимальных объемов закупки, действующих на предприятии;
 – видах транспортных средств для формирования заявок поставщикам с обеспечением максимальной загрузки транспорта.

Одной из основных задач подсистемы логистики является отслеживание прогнозной даты поступления покупных материалов и комплектующих изделий на предприятие. Она влияет на расчет точки наступления дефицита, под которой понимается момент времени, когда прогнозный остаток материала опустится, например, ниже уровня страхового запаса (рис. 2). Смещение указанной прогнозной даты по оси времени вправо может вызвать разрыв в обеспечении производства соответствующими номенклатурными позициями. Если возникает такая ситуация, то возможным способом ее устранения является применение заменяющих материалов из имеющихся в наличии или закупленных с избытком и находящихся в пути на основе ведомости замен.

Номенклатура	Показатель	Источник	01.01	02.01	03.01	04.01	05.01	06.01	07.01	08.01	09.01	10.01	11.01	12.01	13.01	14.01	15.01	16.01
1000 Материал для про...	ОстатокНП	Расчет	50,00	50,00	-100,00	-168,18	-318,18	-518,19	-804,55	-804,55	-804,55	-1 090,91	-485,28	-771,64	-1 058,00	-1 344,37	-1 344,37	-1 344,37
	ДоступноНП	Расчет	50,00	50,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Потребность	Расчет		150,00	68,18	150,00	200,00	286,36			286,36	286,36	286,36	286,36	286,36	286,36	286,36	286,36
	Дефицит	Расчет		100,00	68,18	150,00	200,00	286,36			286,36	286,36	286,36	286,36	286,36	286,36	286,36	286,36
1001 Материал для зан...	ОстатокНП	Расчет	5 000,00	5 000,00	4 862,71	4 725,42	4 588,13	4 450,84	4 313,55	4 313,55	4 313,55	4 203,55	4 093,55	3 983,55	3 983,55	3 983,55	3 983,55	3 983,55
	ДоступноНП	Расчет	5 000,00	5 000,00	4 862,71	4 725,42	4 588,13	4 450,84	4 313,55	4 313,55	4 313,55	4 203,55	4 093,55	3 983,55	3 983,55	3 983,55	3 983,55	3 983,55
	Потребность	Расчет		137,29	137,29	137,29	137,29	137,29			110,00	110,00	110,00					
	Дефицит	Расчет																
1100 Материал БЕЗ ПОС...	ОстатокНП	Расчет	5 000,00	4 862,71	4 725,42	4 588,13	4 450,84	4 313,55	4 313,55	4 203,55	4 093,55	3 983,55	3 983,55	3 983,55	3 983,55	3 983,55	3 983,55	3 983,55
	ДоступноНП	Расчет	801,00	801,00	791,91	782,82	773,73	764,64	755,55	755,55	755,55	746,45	737,36	728,27	719,18	710,09	710,09	710,09
	Потребность	Расчет			9,09	9,09	9,09	9,09	9,09			9,09	9,09	9,09	9,09	9,09	9,09	9,09
	Дефицит	Расчет																
22 Ткань Кронверк 20	ОстатокНП	Расчет	56,40	56,40	13,42	-29,55	-71,07	-71,07	-114,05	-114,05	-114,05	-157,03	-200,00	-242,98	-285,95	-328,93	-328,93	-328,93
	ДоступноНП	Расчет	56,40	56,40	13,42	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Потребность	Расчет			42,98	42,98	41,52		42,98			42,98	42,98	42,98	42,98	42,98	42,98	42,98
	Дефицит	Расчет				29,55	41,52		42,98			42,98	42,98	42,98	42,98	42,98	42,98	42,98
9 Рамы ПМ 1600*2000	ОстатокНП	Расчет	250,00	250,00	204,27	74,27	-145,73	-245,73	-291,45	-291,45	-291,45	-401,45	-197,18	-242,91	-288,64	-334,36	-334,36	-334,36
	ДоступноНП	Расчет	250,00	250,00	204,27	74,27	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Потребность	Расчет			45,73	130,00	220,00	100,00	45,73			110,00	45,73	45,73	45,73	45,73	45,73	45,73
	Дефицит	Расчет																

Рис. 2. Прогноз возникновения дефицита покупных материалов и комплектующих изделий

В ряде случаев, связанных со спецификой свойств номенклатурных позиций, применение теоретических коэффициентов пересчета норм расхода при заменах материалов не имеет практического смысла. Поэтому автоматизированная система должна давать пользователю возможность в конкретных случаях самостоятельно определять соотношения между количеством основного материала и количеством заменяющего его материала, исходя из опыта, знаний и здравого смысла [4]. При работе с допустимыми заменами требуется контролировать, чтобы в результате замены не был нарушен план производства тех изделий, для которых заменяющий материал является основным, а также существовала возможность стоимостного анализа замены, отражающего ее влияние на себестоимость готовой продукции.

Дополнительным аспектом использования ведомости замен является актуализация номенклатуры при расчете потребности в материалах и комплектующих изделиях. Не секрет, что в справочниках материальных норм расхода на многих предприятиях указаны устаревшие, снятые с производства материалы. В таком случае алгоритм планирования в процессе своей работе должен автоматически заменять неактуальные материалы на современные, указанные в качестве обязательной замены.

Необходимой функцией подсистемы логистики является расчет минимальных и максимальных уровней запасов. Расчет может производиться или по статистике потребления материалов и ПКИ в прошедших периодах, например:

$$\text{минимальный запас} = \text{расход за период} / (30 / \text{длительность поставки});$$

максимальный запас = минимальный запас $\cdot k$ \cdot страховой коэффициент,

или по прогнозу их потребления в будущем:

минимальный запас = потребность на период / (30/длительность поставки).

Здесь 30 – нормативное число дней в календарном месяце; k – нормативный коэффициент-мультипликатор, значение которого подбирается эмпирически. Страховой коэффициент также определяется эмпирическим путем для соответствующей номенклатурной группы или позиции.

Средства автоматизации должны иметь опции учета материально-производственных запасов и их движения на производстве в разрезе учетной аналитики (целевого учета), в качестве которой чаще всего выступают заказы потребителей готовой продукции, проекты и их этапы, шифры затрат. Это характерно для позаказного или проектного производства, в его контекст часто попадают несерийные изделия. Данные возможности позволяют осуществлять детальный контроль обеспеченности производства на всех его технологических этапах и в последующем упрощать сбор фактических затрат по соответствующим объектам учета.

Дополнительным плюсом выступает наличие средств управления статусами запасов, отражающих их текущее состояние. Возникновение возмущений в цепочке поставок за счет перевода запасов в определенное состояние (например, брак) приводит к смещению влево по оси времени прогнозной точки наступления дефицита и требует своевременной реакции службы снабжения. Статусы позволяют регулировать вовлечение запасов в процессы расчета обеспеченности и комплектации производственных заданий, а также влиять на процессы запуска партий изделий на цехозаходах (технологических переделах) в процессе диспетчеризации производства.

Подсистема управления логистикой должна обеспечивать возможность профилирования пользователей и их взаимозаменяемость. В результате профилирования каждый специалист службы снабжения работает только со своими профильными номенклатурными позициями. Взаимозаменяемость подразумевает переброску номенклатуры между пользователями, например, на время отпуска одного из них.

Подсистема ведения НСИ и подсистемы логистики являются источником необходимых данных для подсистемы планирования производства изделий и закупок материалов, которая призвана обеспечивать:

- ведение множества вариантов планов продаж в разрезе сбытовых точек и товарных групп;
- непосредственный расчет производственных планов с использованием методики скользящего производственного планирования как минимум на 12 месяцев с детализацией данных по дням на период до 90 дней (в случае необходимости и более), считая от даты начала планового периода;

- несколько вариантов расчета производственных планов для каждого варианта плана продаж (с учетом или без учета загрузки мощностей, нормативов времени опережения запуска заказов в производство и иных параметров, влияющих на результаты планирования);

- варьирование объемов доступных мощностей по периодам планирования. Для этого информационная структура системы должна давать возможность ведения:

- 1) вариантов графиков рабочего времени с учетом плановых простоев технологического оборудования на ремонт и техническое обслуживание, графиков отпусков основного производственного персонала;

- 2) вариантов графиков ввода-вывода производственных мощностей в эксплуатацию и из эксплуатации, найма или сокращения рабочей силы;

- 3) эмпирических коэффициентов использования производственных мощностей и рабочего времени трудовых ресурсов по плановым периодам, рассчитанных на основе статистических данных о внеплановых простоях оборудования из-за аварий, потерях по больничным листам и т. п.;

- оперативный пересчет всей совокупности производственных планов как минимум один раз в несколько дней (желательно ежедневно);

- управление запасами сырья, материалов и комплектующих изделий как по прогнозной точке наступления дефицита (обеспечение основного производства), так и по факту снижения уровня запаса ниже критического уровня (обеспечение запчастями и другими вспомогательными материалами);

- планирование закупок комплектов;
 - получение наглядной аналитической отчетности и сравнительный анализ различных вариантов расчета на основе OLAP(Online Analytical Processing)-технологии.
- Результатами работы, построенной на вышеописанных принципах системы, являются:
- оценка реалистичности планов продаж с отражением возможных отклонений;
 - графики производства и отгрузки товаров в сбытовые точки, сбалансированные с доступными материалами и мощностями (рис. 3);

Наимен. Тип	Январь 2011	Февраль 2011	Март 2011	Апрель 2011	Май 2011	Июнь 2011	Июль 2011	Август...
И 1.0 Издел... ОстатокНП	18 950,00	429,00	242,00	167,00	1 036,00	277,00	1 794,00	1
РасчетныйВыпуск	32 725,00	30 730,00	25 382,00	43 605,00	36 167,00	12 691,00		
ОтгрузкаВСС	51 246,00	30 917,00	25 457,00	42 736,00	36 926,00	11 174,00		
ОстатокКП	429,00	242,00	167,00	1 036,00	277,00	1 794,00	1 794,00	1
И 2.0 Издел... ОстатокНП	0,00	0,00	0,00	0,00	0,00	0,00	873,00	
РасчетныйВыпуск						35 353,00	62 700,00	64
ОтгрузкаВСС						34 480,00	62 688,00	65
ОстатокКП	0,00	0,00	0,00	0,00	0,00	873,00	885,00	
И 3.0 Издел... ОстатокНП	700,00	63,00	63,00	1 442,00	31,00	51,00	17,00	
РасчетныйВыпуск	1 381,00		1 379,00	1 383,00	1 381,00	2 762,00	1 381,00	
ОтгрузкаВСС	2 018,00			2 794,00	1 361,00	2 796,00	1 291,00	
ОстатокКП	63,00	63,00	1 442,00	31,00	51,00	17,00	107,00	
И 4.0 Издел... ОстатокНП	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
ОстатокКП	0,00	0,00	0,00	0,00	0,00	0,00	0,00	

Рис. 3. График выпуска и отгрузки продукции

– рекомендации по срокам размещения и объемам закупок заказов на материалы и комплектующие изделия у поставщиков с выбором оптимального маршрута их доставки, которые предупреждают сбои в обеспечении основного производства материалами и ПККИ. Расчет объемов закупок производится по правилам [5], указанным для номенклатурных позиций: строго под потребность (с учетом или без учета складских остатков); на пополнение запаса до максимального уровня; на обеспечение потребности производства на заданный временной буфер (например, на две недели). Срок размещения рассчитывается исходя из нормативов длительности этапов жизненного цикла заказов на закупку (согласование с поставщиком, производственный цикл поставщика (с момента согласования заявки до момента отгрузки), доставка, нахождение на входном контроле и т. п.). В ряде случаев для сокращения объема обработки и администрирования нормативно-справочных данных можно пойти на группировку номенклатурных позиций по типовой длительности циклов закупки (например, 30, 60 или 90 дней). Под длительностью цикла закупки тогда понимается период времени с момента отправки заявки поставщику на согласование до момента поступления заказанных материалов на склад предприятия без деления данного периода на промежуточные этапы;

- тренды потребности в материалах и комплектующих изделий на основе статистики, благодаря которым сотрудники службы снабжения могут в ручном режиме корректировать рекомендованные системой объемы закупок;
- оперативный контроль размещения и состояния исполнения заказов на закупку;
- упреждающая индикация возможности возникновения дефицита ПККИ из-за окончания их сроков годности, гарантийных сроков хранения или истечения сроков действия сертификатов поставщиков.

Заключение

Предложенный в работе подход прошел практическую проверку при создании автоматизированной системы планирования производства Галактика АММ (Advanced Manufacture Management) для предприятий, производящих серийную продукцию с высокой долей в ее со-

ставе ПКИ. Для предприятий характерны короткие (до семи дней) технологические циклы производства, длительные сроки поставки комплектующих изделий (до трех месяцев) и сроки доставки готовой продукции в сбытовые точки (до 30 дней). Номенклатурный справочник покупных материалов и комплектующих может включать десятки тысяч номенклатурных позиций.

Список литературы

1. Питеркин, С.В. Точно вовремя для России. Практика применения ERP-систем / С.В. Питеркин, Н.А. Оладов, Д.В. Исаев. – М. : Альпина паблишер, 2003. – 368 с.
2. Линдерс, М. Управление снабжением и запасами. Логистика / М. Линдерс, Х. Фирон ; пер. с англ. – СПб. : ООО «Полиграфуслуги», 2006. – 768 с.
3. Современные методы организации, управления и технологии бережливого производства : учеб. пособие / Р.Х. Зарипов [и др.]. – Казань, 2010. – 304 с.
4. Шатохин, И.В. Бизнес, точный как часы / И.В. Шатохин // Белорусы и рынок. – 2011. – № 48(983). – С. 14.
5. Управленческий учет снабженческо-заготовительной деятельности / В.Э. Керимов [и др.] ; под ред. В.Э. Керимова. – М. : Экзамен, 2002. – 128 с.

Поступила 10.10.2016

УП «Топ Софт»,
Минск, ул. Сурганова, 28В
e-mail: IVS@galaktika.by

I.V. Shatokhin

INFORMATION SYSTEM OF ENTERPRISE SUPPLY CHAIN MANAGEMENT

The supply chain management method for manufacture scheduling and planning based on APS-algorithm is offered. It . Logistics solutions and inventory management problems are discussed. The structure of automated control system of enterprise is designed and its tasks are described.

УДК 621.396

М.А. Смольников¹, Ю.А. Скудняков²

ИСПОЛЬЗОВАНИЕ АГРЕГАЦИИ ДАННЫХ В МОБИЛЬНЫХ СЕНСОРНЫХ СЕТЯХ

Проводится сравнение традиционных алгоритмов сетевой маршрутизации с алгоритмом, выполняющим перенаправление сообщений с учетом их содержимого. Исследуется влияние расположения источника данных и получателя на энергозатраты, задержки и надежность передачи данных. Доказывается, что подобная маршрутизация может увеличить эффективность работы сенсорных сетей в широком диапазоне условий.

Введение

Сенсорные сети обладают рядом отличий от традиционных сетей передачи данных, устройства которых не имеют надежного источника питания и постоянно передают небольшие объемы данных от множества точек, распределенных географически, к единственному серверу. В связи с этим применение стандартных схем маршрутизации в сенсорных сетях является недостаточно эффективным.

Мобильная беспроводная сенсорная сеть – это сеть, состоящая из множества недорогих узлов, каждый из которых имеет собственный микропроцессор, набор датчиков и сетевой интерфейс. Устройства максимально автономны и могут работать без вмешательства оператора. Область применения беспроводных сенсорных сетей достаточно широка: начиная от простого отслеживания транспорта и заканчивая мониторингом окружающей среды. Основа аппаратной части – недорогие процессоры, обладающие малой производительностью (например, базирующиеся на ядре AVR).

В настоящее время уже существуют как проприетарные, так и свободно распространяемые программные решения для осуществления сбора и обработки сигналов, поступающих с подключенных к микроконтроллеру сенсоров, однако коммуникация между устройствами до сих пор ведется преимущественно с использованием стандартных сетевых протоколов (например, IEEE802.11).

Разработка новых решений для сбора данных с сенсорной сети часто базируется на архитектуре MANET (MobileAd-HocNetwork) – беспроводной децентрализованной самоорганизующейся сети. Каждое устройство сети может независимо передвигаться и, как следствие, создавать и разрывать соединения с соседями. Кроме того, к устройствам предъявляется требование возможности функционирования в качестве роутера. В настоящее время на базе этой технологии реализовано множество известных проектов, например SARTRE [1], компанией Volvo разработана система автопилота для автомобилей, движущихся в колонне. Сети для коммуникации транспортных средств тем не менее имеют определенную специфику (повышенные требования к безопасности и отказоустойчивости), поэтому их чаще выделяют в отдельный подвид – VANET [2].

Особенностью же мобильных сенсорных сетей (МСС), в частности сетей, которые измеряют характеристики окружающей среды (например, барометрические, сейсмические, температурные и др.), является то, что с большой вероятностью показания соседних датчиков, а следовательно, и данных, которые они передают, будут одинаковы или похожи друг на друга. Второй ключевой особенностью МСС следует назвать передачу данных только в одном направлении: от множества источников к одному приемнику. Третья важная деталь: собственно сенсоры могут находиться в таких местах, где обслуживание их человеком может быть затруднено или вовсе невозможно. Несмотря на то что классическая MANET способна работать в обозначенных условиях, ее эффективность может быть увеличена за счет агрегации данных и отсутствия поддержки коммуникации с произвольным узлом.

На рис. 1 изображены потоки данных в обычной сети Ad-Нос (слева) и сети с применением агрегации (справа). Красным цветом выделены узлы – источники данных, синим цветом – получатель, а черным – промежуточные узлы, выполняющие роль маршрутизаторов. Стрелками обозначены потоки данных, цвет стрелок условно показывает различия в значении регистрируемого параметра. За счет применения агрегирующей функции на ретрансляторах изображенной сети существует возможность уменьшить число передач с 8 до 5.

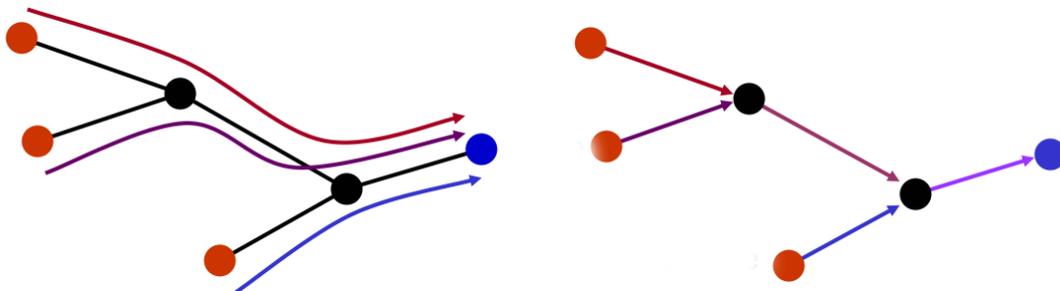


Рис. 1. Сравнение потоков данных в обычной сети Ad-Нос и сети с применением агрегации

Агрегация данных – это процесс комбинирования данных из множества источников, который может быть реализован множеством способов. Простейшим способом является удаление повторяющихся данных, агрегация может быть представлена функцией (например, поиска минимального, максимального или среднего значения) с количеством аргументов более одного. Для целей моделирования сделано допущение: функция-агрегатор будет выдавать один пакет данных при любом количестве входных аргументов.

Алгоритм маршрутизации с использованием агрегации данных в мобильных сенсорных сетях

В простейшем случае процесс агрегации может быть рассмотрен как широкое вещание наоборот: вместо того чтобы отправлять один пакет от источника множеству получателей, множество передатчиков отправляет один и тот же пакет данных приемнику. Допустим, есть k источников, обозначенных от S_1 до S_n , а также приемник, представленный как D . Пусть граф сети $G = (V, E)$ состоит из узлов V , связанных между собой ребрами E . Нахождение оптимального числа передач, требуемых для достижения данными получателя, сводится к решению задачи Штейнера [3]. Однако стоит отметить, что существующие алгоритмы решения этой задачи не обладают высокой эффективностью. Так, алгоритм Крускала [4], считающийся одним из оптимальных средств решения задачи, требует выполнения перечня достаточно сложных операций. Наиболее затратные из них – сортировка ребер и представление связей в виде системы непересекающихся множеств. Время работы алгоритма Крускала зависит от числа ребер и может быть определено как $O(E \cdot \log E)$. Возвращаясь к основным требованиям, предъявляемым к сети (необходимость самоорганизации и использование низкопроизводительных компонентов), следует отметить, что процесс нахождения маршрутов на больших сетях может занимать длительное время.

Рассмотрим три основных вида генерации деревьев агрегации данных:

ленивая – агрегацию выполняет только ближайший к источнику узел, а прочие просто отправляют данные прямо к источнику;

быстрая – каждый источник стремится отправить данные по наиболее короткому пути;

жадная, – напротив, нацелена на достижение максимальной агрегации: только наиболее близкий к серверу источник имеет прямую связь с ним, остальные присоединяются к самым близким узлам.

Для поиска оптимальной схемы рассмотрим следующие показатели:

– энергозатраты. Полагая, что операция агрегации «дешевле», чем операция передачи данных, число передач может быть уменьшено с целью снижения энергопотребления;

– задержки. Очевидно, что при использовании агрегации, перед тем как произвести отправку данных далее, необходимо сначала дождаться информации от всех узлов, а затем вы-

полнить над ней определенные действия. Следовательно, время доставки пакетов серверу будет значительно увеличено;

– робастность. Подключение новых узлов к существующей сети с агрегацией – значительно более простая операция, так как не всегда требует пересчета сети целиком.

Рассчитаем эффективность алгоритма маршрутизации с использованием агрегации данных в сравнении с простой маршрутизацией. Пусть d_i – наиболее короткое расстояние от источника S_i до потребителя данных на графе. Для каждой передачи число перенаправлений, требуемых для оптимального классического адресного протокола, в этом случае можно найти как сумму расстояний между перенаправляющими узлами:

$$N_A = d_1 + d_2 + \dots + d_{n-1} + d_n = \sum_{i=1}^n d_i.$$

В алгоритме, использующем агрегацию данных, число передач N_D всегда будет меньшим либо равным N_A . Диаметр X набора вершин S на графе G – наиболее длинный из попарно кратчайших путей между всеми вершинами. Если диаметр X больше либо равен единице, то общее число передач, требуемое для оптимального протокола, будет удовлетворять условию неравенства

$$\begin{cases} N_D \leq (n-1)X + \min(d_i), \\ N_D \geq \min(d_i) + (n-1). \end{cases}$$

Верхняя граница в этом случае определяется исходя из того, что наибольшее число ребер в графе с $(n-1)$ отправителями (оставшаяся вершина является получателем данных) не может быть больше произведения диаметра X на число передающих вершин плюс расстояние от источника до ближайшей передающей вершины. Нижняя граница – частный случай дерева Штейнера для $X=1$, когда все узлы передают данные непосредственно агрегатору.

Как следствие, если $X < \min(d_i)$, то и $N_D < N_A$. Значит, протокол маршрутизации, использующий агрегацию, будет работать эффективнее, чем обычное перенаправление данных:

$$N_D < (n-1)X + \min(d_i) < n\min(d_i) \Rightarrow N_D < \text{sum}(d_i) = N_A.$$

Относительную эффективность можно определить следующим образом:

$$E = \frac{(N_A - N_D)}{N_A}.$$

Раскрыв значения N_A и N_D , получим границы значения эффективности как диапазон от 0 до 1:

$$\begin{cases} E \geq 1 - \frac{((n-1)X + \min(d_i))}{\text{sum}(d_i)}, \\ E \leq 1 - \frac{(\min(d_i) + n - 1)}{\text{sum}(d_i)}. \end{cases}$$

Расположив все узлы на равном удалении от получателя: $\min(d_i) = \max(d_i) = d$, имеем

$$1 - \frac{(n-1)X + d}{nd} \leq E \leq 1 - \frac{d + n - 1}{nd}.$$

Если предположить, что значения X и n неизменны и значительно меньше, чем d , то при увеличении количества перенаправлений d эффективность будет зависеть от общего числа узлов:

$$\lim_{d \rightarrow \infty} E = 1 - 1/n.$$

Очевидно, что в случае, если дистанция между сервером и остальными узлами достаточно велика и сами узлы образуют между собой кластер, протокол с поддержкой агрегации дает большой прирост эффективности: 50 % при двух узлах в сети и 90 % при 10 узлах.

Для экспериментального подтверждения полученных результатов построена модель, в которой на плоскости случайным образом размещены передатчики. В модели учитываются два фактора: радиус, внутри которого произошло событие (расстояние, в рамках которого все датчики будут регистрировать и передавать одинаковое значение какой-либо величины), и радиус действия радиомодуля узлов. Моделирование проводится в приложении Omnet++, число узлов 100 (99 узлов и 1 сервер). Выбрана предустановленная модель в приложении симуляции модели Ad-Нос, модифицированная дополнительным условием: если в узел-маршрутизатор приходят два пакета одинаковых данных, следует отправить только один пакет данных на сервер.

Число перенаправлений в моделируемой сети с использованием агрегации зависит от двух параметров: радиуса связи между узлами и радиуса регистрации события (рис. 2).

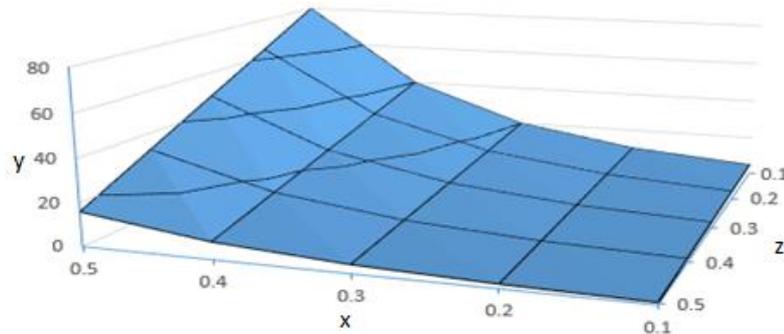


Рис. 2. Зависимость числа перенаправлений одного пакета данных от радиуса связи узлов (ось X) и размера события (ось Z)

Для сравнения эффективности сети, использующей агрегацию, с традиционной сетью Ad-Нос, информация об одинаковых событиях одного и того же размера передана через обе сети (рис. 3). Измерено число перенаправлений (вертикальная ось) в зависимости от размера события (горизонтальная ось).

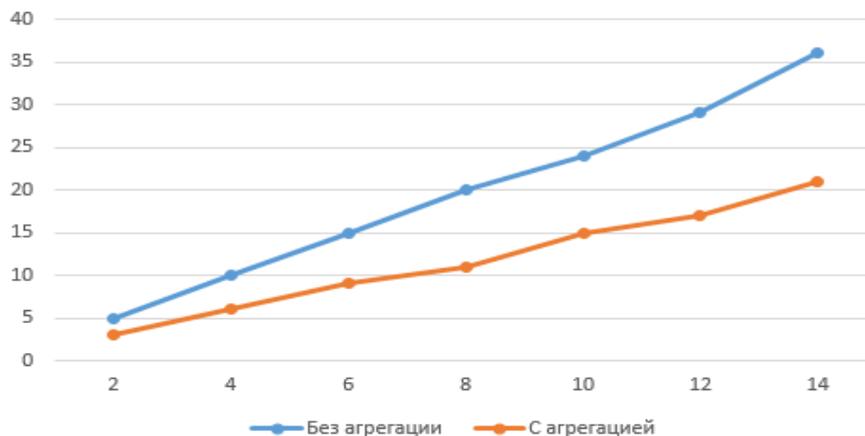


Рис. 3. Сравнение числа перенаправлений в схеме маршрутизации без агрегации и с агрегацией данных

Результаты сравнения показали, что в определенных условиях, когда одно и то же событие регистрируется значительным числом датчиков, т. е. радиус связи значительно меньше, чем радиус события, число перенаправлений может быть достаточно малым (см. рис. 2). В то же время при отсутствии агрегации такой параметр, как радиус события, не влияет на уменьшение числа передач и будет максимальным для данной дальности связи (см. рис. 3).

Принимая также во внимание, что увеличение дальности связи влечет за собой экспоненциальный рост потребляемой мощности (возможно использование направленных антенн, уменьшающих данный эффект) и каждая передача данных требует значительных затрат энергии как передающего, так и принимающего узлов, максимальная агрегация данных с целью уменьшения объема передачи и времени на согласование связи, требующая лишь вычислительной мощности микроконтроллера, при достаточной эффективности может значительно продлить время работы, разгрузить частоту, на которой происходит передача, а следовательно, увеличить максимальную плотность сети.

Кроме рассмотренных методов агрегации стоит обратить внимание на алгоритмы сжатия данных, однако это направление требует дальнейших исследований, так как для использования более эффективных алгоритмов необходимо больше процессорного времени, что вызовет увеличение энергопотребления каждым перенаправляющим узлом.

Заключение

В работе осуществлен анализ существующих МСС с точки зрения их назначения, функциональных возможностей, практической реализации, достоинств и недостатков. На основе результатов проведенного исследования предложен алгоритм оптимизации маршрутизации в МСС с агрегацией данных, использование которого дает возможность сократить объем передаваемой информации при сохранении надежности связей в сетях. Применение разработанного алгоритма оптимизации маршрутизации позволяет сенсорной сети функционировать с наименьшими энергозатратами по сравнению с существующими сетями.

Список литературы

1. Chapman & Hall CRC Computer and Information science series Distributed Sensor Networks S. Sitharma Iyengar and Richard R. Brooks [Electronic resource]. – 2011. – Mode of access : [http:// bit.ly/29c3Gw6](http://bit.ly/29c3Gw6). – Date of access : 24.05.2016.
2. Галкин, П.В. Анализ энергопотребления узлов беспроводных сенсорных сетей / П.В. Галкин // ScienceRise. – 2014. – № 2. – С. 20–22.
3. Восков, Л.С. Повышение качества обслуживания в беспроводных стационарных сенсорных сетях с автономными источниками питания / Л.С. Восков, М.М. Комаров // Качество. Инновации. Образование. – 2012. – № 1. – С. 80–88.
4. Dynamic Source Routing in Ad-Hoc Wireless Networking / D.B. Johnson [et al.]. – Pittsburgh : Mobile Computing, 1996. – 25 p.

Поступила 12.10.2016

¹Белорусский национальный
технический университет,
Минск, пр. Независимости, 65
e-mail: me@esight.ru

²Белорусский государственный университет
информатики и радиоэлектроники,
Минск, П. Бровки, 6
e-mail: juri_alex@tut.by

M.A. Smolnikov, Y.A. Skudnyakov

THE USE OF DATA AGGREGATION IN MOBILE SENSOR NETWORKS

In the article the comparison of traditional routing algorithms with content based routing algorithms is performed. Investigation of source and sink locations influence on energy consumption, delays and data transfer reliability is made as well. It is proved, that this kind of routing may increase the efficiency of sensor networks operating in a large range of conditions.

ПРАВИЛА ДЛЯ АВТОРОВ

1. Статьи принимаются в редакцию через электронную систему подачи по адресу <http://jinfo.bas-net.by> в формате файлов текстовых редакторов Microsoft Word 97 и Word 2000 для Windows. Основной текст статьи набирается с переносами шрифтом Times New Roman 11 пт, интервал между строками – одинарный, абзацный отступ 1 см, поля по 2,5 см со всех сторон.

2. Статья должна иметь индекс УДК (универсальная десятичная классификация).

3. Название статьи, фамилии всех авторов и аннотация должны быть переведены на английский язык. Для каждого из авторов приводится развернутое название учреждения с полным почтовым адресом, а также номер телефона и электронный адрес (e-mail) для связи с редакцией.

4. Формулы, иллюстрации, таблицы, встречающиеся в статье, должны быть пронумерованы в соответствии с порядком цитирования в тексте. Ссылки на рисунки и таблицы в тексте обязательны. Необходимо избегать повторения одних и тех же данных в таблицах, графиках и тексте статьи.

Рисунки должны быть выполнены с хорошим разрешением в масштабе, позволяющем четко различать надписи и обозначения. Подрисовочные подписи с расшифровкой всех позиций, представленных на рисунке, набираются шрифтом гарнитуры основного текста, размер символов 9 пт. Цветные иллюстрации печатаются только в том случае, когда это необходимо для понимания излагаемого материала.

5. Набор формул выполняется в формульных редакторах Microsoft Equation или Math Type и должен быть единообразным по применению шрифтов и знаков по всей статье.

Прямо () набираются: греческие и русские буквы; математические символы (\sin , \lg , ∞); символы химических элементов (C, Cl, CHCl_3); цифры (римские и арабские); векторы; индексы (верхние и нижние), являющиеся сокращениями слов.

Курсивом (–) набираются: латинские буквы – переменные, символы физических величин (в том числе и в индексе).

6. Сокращения в тексте статьи (за исключением единиц измерения) могут быть использованы только после упоминания полного термина. Единицы измерения физических величин следует приводить в Международной системе СИ.

7. Литература приводится автором общим списком в конце статьи. Ссылки на литературу в тексте идут по порядку и обозначаются цифрой в квадратных скобках. Ссылаться на неопубликованные работы не допускается. С примерами оформления библиографического описания в списке литературы можно ознакомиться в приложении 2 к *Инструкции по оформлению диссертации, автореферата и публикаций по теме диссертации* на сайте Высшей аттестационной комиссии Республики Беларусь <http://vak.org.by>.

8. Поступившие в редакцию статьи направляются на рецензирование специалистам. Основным критерием целесообразности публикации является новизна и информативность статьи. Если по рекомендациям рецензента статья возвращается автору на доработку, а переработанная рукопись вновь рассматривается редколлегией, датой поступления считается день получения редакцией ее окончательного варианта. Статьи не по профилю журнала возвращаются авторам после заключения редколлегии.

9. Статьи, направляемые на доработку, должны быть возвращены в исправленном виде с ответами на все вопросы.

10. Редакция журнала предоставляет возможность первоочередного опубликования статей, представленных лицами, которые осуществляют послевузовское обучение (аспирантура, докторантура, соискательство) в год завершения обучения.

11. Авторы несут ответственность за направление в редакцию статей, уже опубликованных ранее, или статей, принятых к публикации другими изданиями.

12. Редакция оставляет за собой право на редакционные изменения, не искажающие основное содержание статьи.

Журнал «Информатика» включен Высшей аттестационной комиссией Республики Беларусь в список научных изданий для опубликования результатов диссертационных исследований.

Индексы

00827

для индивидуальных
подписчиков

008272

для предприятий и
организаций