

6. *Кутенов А.М., Бондарева Т.Н., Беренгартен М.Г.* Общая химическая технология. – М.: Высш. шк. 1990. – 520 с.
7. *Осинов А.И., Панченко В.Я.* Тепловые эффекты при взаимодействии лазерного излучения с молекулярными газами. – М.: МГУ. 1983. – 117 с.
8. *Леше А.* Физика молекул. М. Мир.. 1987 – 232 с.
9. *Тауер Дж., Нехвытал Э.* Орбитальная Теория в контурных диаграммах. – М.: Мир. 1988.- с.124.
10. *Сікора Л.С.* Лазерні інформаційно-вимірвальні системи для управління технологічними процесами. – Львів.: Каменяр. 1988. – 445с.
11. *Пасинский А.Г.* Биофизическая химия – М.: Высш. шк. 1968. - 43 с.
12. *Степанов Н.Ф.* Квантовая механика и квантовая физика. – М.: Мир. 2001. - 519 с.

*Поступила 12.10.2015р.*

УДК 004.032.26

І.Г. Цмоць<sup>1</sup>, Т.В. Теслюк<sup>1</sup>, І.В. Ігнатєв<sup>2</sup>

<sup>1</sup> Національний університет “Львівська політехніка”

<sup>2</sup>Тернопільський національний економічний університет

## **ПРОЕКТУВАННЯ СПЕЦІАЛІЗОВАНИХ КОМПОНЕНТ НЕЙРО-ОРІЄНТОВАНИХ КОМП'ЮТЕРНИХ СИСТЕМ**

Сформований операційний базис спеціалізованих компонентів нейро-орієнтованої комп'ютерної систем, визначені вимоги до алгоритмів функціонування та структури спеціалізованих компонентів, вибрано принципів проектування та розроблено метод синтезу спеціалізованих компонентів нейро-орієнтованої комп'ютерної систем, який враховує особливості елементної бази, вимоги конкретного застосування та забезпечує обробку даних в реальному часі.

**Ключові слова:** нейро-орієнтована комп'ютерна системи, спеціалізована компонента, синтез, структура, НВІС-реалізація.

Formed operational basis of the specialized components of the neuro-oriented computer systems. Defined requirements to algorithms of functioning and structure of specialized components. Selected principles of design and developed synthesis method of specialized components of neuro-oriented computer systems, which considers particular features of the hardware components, requirements of particular application and provides real time data processing.

**Key words:** neuro-oriented computer systems, specialized component, synthesis, structure, VLSI implementation.

## **Постановка проблеми**

Сучасний етап розвитку нейротехнологій характеризується розширенням галузей їхнього застосування, більшість з яких потребують опрацювання інтенсивних потоків даних у реальному часі на апаратних засобах, що задовольняють обмеження відносно габаритів, маси, енергоспоживання тощо. До таких галузей належить промисловість (управління технологічними процесами та складними об'єктами), енергетика (оптимізація навантаження в електромережах), військова (технічний зір, управління рухом мобільного робота), автомобільна (управління рухом і двигуном), медична (діагностика захворювань) і приладобудівна (розпізнавання образів і оптимізація управління). Для реалізації штучних нейронних мереж у більшості випадків використовують комп'ютерні системи з структурною організацією універсального типу, яка функціонально та структурно є надлишковою, не враховує вимог конкретних застосувань щодо продуктивності, габаритів, споживаної потужності та має невисоку ефективність використання обладнання.

Створення комп'ютерних систем, орієнтованих на ефективну реалізацію у реальному часі нейромережових алгоритмів, досягається доповненням їх спеціалізованими компонентами, які реалізують базові нейрооперації, нейроелементи, шари нейроелементів та нейромережі. Проектування таких спеціалізованих компонентів потребує широкого використання сучасної елементної бази, розроблення алгоритмів і спеціалізованих НВІС-структур. Режим реального часу та НВІС-реалізація спеціалізованих компонентів забезпечується розпаралелюванням і конвеєризацією процесів обчислень, апаратним відображенням структури нейромережових алгоритмів у архітектуру, яка адаптована до інтенсивності надходження потоків даних.

З огляду на це особливої актуальності набуває завдання проектування спеціалізованих компонентів і доповнення ними комп'ютерних систем для забезпечення нейромережевого опрацювання даних у реальному часі.

## **Аналіз останніх досліджень та публікацій**

Аналіз підходів до реалізації нейромережових алгоритмів [1-11] показав, що із множини існуючих підходів можна виділити такі:

- на основі паралельних універсальних і функціонально-орієнтованих процесорів шляхом розробки спеціалізованого програмного забезпечення;
- нейро-орієнтований, який передбачає використання обчислювального ядра, орієнтованого на задачі нейромережевої обробки, доповненого спеціалізованими апаратно-програмними компонентами, які реалізують базові операції нейромережових алгоритмів;
- апаратно-спеціалізований, архітектура та організація обчислювального процесу в якому відображає структуру нейромережевого алгоритму розв'язання задачі.

Перший підхід є доступним для широкого кола користувачів. Істотною його перевагою є можливість використання раніше розроблених програм. Недоліками такого підходу є невисока швидкодія, функціональна і структурна надлишковість комп'ютерних засобів.

Другий підхід передбачає поєднання універсальних і спеціальних засобів. Процес взаємопроникнення універсального та спеціального, програмного і апаратного забезпечує високу ефективність використання обладнання та опрацювання у реальному часі потоків даних за нейромережевими алгоритмами. При такому підході розробка нейроорієнтованої комп'ютерної систем з заданими технічними параметрами зводиться до доповнення обчислювального ядра необхідними спеціалізованими апаратно-програмними компонентами.

Апаратно-спеціалізований підхід орієнтований на опрацювання інтенсивних потоків даних за складними алгоритмами. Такий підхід має такі недоліки:

- апаратні спеціалізовані нейрокомп'ютерні системи створюються для вирішення тільки конкретних задач;
- складність процедури навчання (самонавчання) системи;
- необхідні великі затрати обладнання на їх реалізацію;
- апаратні спеціалізовані нейрокомп'ютерні системи реального часу не враховують вимоги конкретних застосувань щодо габаритів і споживаної потужності;
- велика вартість і час розробки апаратних спеціалізованих нейрокомп'ютерних систем реального часу.

З проведеного аналізу видно, що перспективним підходом до реалізації нейромережових алгоритмів у реальному часі з заданими технічними характеристиками є нейро-орієнтований підхід, який передбачає створення на базі обчислювального ядра та спеціалізованих апаратно-програмних компонентів нейро-орієнтованої комп'ютерної системи. Використання універсального та спеціального підходів, програмних і апаратних засобів забезпечує розширення галузей застосування та високу ефективністю використання обладнання.

*Метою роботи* є формування операційного базису, вимог до алгоритмів функціонування та структури спеціалізованих компонентів, вибір принципів проектування та розроблення методу синтезу спеціалізованих компонентів нейро-орієнтованої комп'ютерної системи.

## **Основна частина**

### ***Компоненти нейро-орієнтованої комп'ютерної системи.***

Нейроорієнтована комп'ютерна системи складається з двох частин: постійної - універсального обчислювального ядра та змінної - спеціалізованих компонентів, які реалізують базові операції нейроалгоритмів [1-3]. Основою обчислювального ядра нейро-орієнтованої комп'ютерної системи є

нейропроцесор (нейрочіп), який має набір команд добре пристосованих для виконання базових операцій нейроалгоритмів, а також повний набір команд загального призначення. Переважна більшість нейалгоритмів зводиться до виконання обмеженого набору базових операцій типу “додавання – множення”

Основними перевагами сучасних нейрочіпів є: відносно більша швидкодія (порівняно з CPU), полегшена реалізація зв'язків «всі зі всіма» (для розробника нейромереж), низьке споживання електроенергії та відносно доступна ціна. Недоліками таких нейрочіпів є: велика структурна складність і низька надійність системи; велика складність ефективною реалізації процедури навчання, самонавчання, самоорганізації; значне збільшення споживаної потужності та втрата швидкодії під час збільшення ступеню інтеграції нейрочіпів; жорстко задана наперед топологія [11].

Основні характеристики нейро-орієнтованої комп'ютерної системи визначаються особливостями архітектури і технічними характеристиками нейропроцесора а саме: довжина інформаційного слова; число основних команд і час їх виконання; ємність пам'яті, що може адресуватися; ємності внутрішньокристалічної пам'яті даних і програм та кількість внутрішніх регістрів.

Структура нейро-орієнтованої комп'ютерної системи, тобто склад змінної частини (кількість і види спеціалізованих компонентів), визначається вимогами конкретних застосувань. Підвищення швидкодії нейро-орієнтованої комп'ютерної системи досягається включенням в склад комп'ютерної системи спеціалізованих компонент, які апаратно реалізує найбільш часомісткі нейроалгоритми та базові нейромережеві операції. Нейромережний операційний базис спеціалізованих компонентів складається із трьох груп базових операцій [1]: попередньої обробки, процесорних операцій та елементарних функцій.

В першу групу входять такі операції: обчислення максимальних і мінімальних чисел, ділення, добування квадратного кореня, обчислення суми квадратів різниць.

Друга група процесорних операцій складається із таких операцій: обчислення скалярного добутку та групового підсумовування.

Третя група складається із операцій: обчислення уніполярної та біполярної порогової функції, гіперболічного тангенса та синусоїдальних функцій з насиченням.

**Вибір форми відображення нейроалгоритмів для розробки апаратних компонент.** Для апаратної реалізації нейроалгоритмів є потреба у їх просторово-часовому відображенні на рівні арифметичних і базових операцій [1,3]. Таке відображення алгоритму повинно забезпечувати виявлення всіх форм паралелізму та знаходження необхідних просторово-часових рішень для проектування спеціалізованих компонентів.

Забезпечити просторово-часове відображення алгоритмів з виявленням всіх форм паралелізму дозволяє ярусно-паралельна форма (ЯПФ), тобто

потоків граф [3]. При такій формі подання алгоритму здійснюється розподіл всіх його функціональних операторів  $\Phi_i$  за ярусами таким чином, що в  $j$ -му ярусі розміщені функціональні оператори, які залежать хоча б від одного функціонального оператора  $(j-1)$ -го ярусу і не залежать від операторів наступних ярусів. Всі функціональні оператори одного ярусу виконуються незалежно один від одного.

В залежності від засобів реалізації, відображення потоків графів алгоритмів може здійснюватися з різною ступенем деталізації. Для реалізацій у вигляді надвеликих інтегральних схем (НВІС) використовується відображення алгоритмів на рівні арифметичних операцій.

Синтез високоефективних спеціалізованих компонентів для нейроорієнтованих комп'ютерних систем реального часу вимагає розробки відповідних алгоритмів і узгоджених потоків графів. Процес розробки узгоджених потоків графів можна розбити на наступні чотири етапи:

- декомпозиція алгоритму розв'язання задачі;
- проектування комунікацій (обмін даними) між функціональними операторами;
- укрупнення функціональних операторів;
- планування обчислень.

На етапі декомпозиції алгоритм розв'язання задачі  $\Phi$  розбивається на функціональні оператори  $\Phi_i$ , між якими устанавлюються зв'язки, що відповідають даному алгоритму.

На етапі проектування комунікацій визначається структура каналів обміну даними між функціональними операторами  $\Phi_i$ . Для чого виконується перехід від граф схеми алгоритму до потоків графу, в якому здійснюється просторово-часове розміщення і закріплення функціональних операторів  $\Phi_i$  за ярусами. Структура зв'язків у потоковому графі між функціональними операторами  $\Phi_{jk}$  сусідніх ярусів визначає кількість каналів надходження даних і структуру з'єднань.

Отримані потокові графи забезпечують оцінювання обчислювальної здатності апаратної компоненти. Для випадку коли обчислювальна здатність апаратної компоненти є значно більшою від необхідної переходять до наступних етапів розробки узгодженого потоковий графу.

На етапі укрупнення здійснюється об'єднання функціональних операторів  $\Phi_{jk}$  і каналів передачі даних як у ярусах потоковий графа, так і між ярусами. Граф алгоритму, який отримаємо у результаті такого об'єднання будемо називати конкретизованим потоковий графом. Етап укрупнення тісно пов'язаний етапом планування, на якому після об'єднання функціональних операторів для збереження інформації про структуру потоковий графа алгоритму здійснюється планування обчислень, визначаються величин затримок і перестановки даних. Для відтворення обчислень у кожний ярус конкретизованого графа вводяться оператори управління затримки та перестановки даних.

**Метод синтезу спеціалізованих компонентів.** Метою проектування спеціалізованих компонентів нейро-орієнтованих комп'ютерних систем реального часу є отримання модульної та регулярної структури орієнтованої на НВІС-технологію [1,3]. Вихідною інформацією для синтезу спеціалізованих компонентів є:

- тип макрооперації (попереднє опрацювання, нейроелемент, шар нейроелементів, функція активації), яку необхідно реалізувати;
- обсяг вхідних даних і інтенсивність їх надходження;
- вимоги до інтерфейсу;
- техніко-економічні вимоги і обмеження.

Задача проектування спеціалізованих компонентів зводиться до забезпечення їхнього функціонування в реальному часі при мінімальних апаратних затратах [3]. Перехід від алгоритму функціонування спеціалізованої компоненти до її структури формально зводиться до мінімізації апаратних затрат:

$$W_{СК} = W_{ПО} + W_{HE} + W_{П} + W_{БУ},$$

де  $W_{ПО}$ ,  $W_{HE}$ ,  $W_{П}$ ,  $W_{БУ}$  – витрати обладнання на реалізацію відповідно на попередню обробку, на реалізацію нейроелементів, пам'ять та блок управління.

В загальному задачу проектування спеціалізованих компонентів нейро-орієнтованих комп'ютерних систем реального часу можна сформулювати так:

- розробити алгоритм функціонування спеціалізованої компоненти;
- представити алгоритм функціонування спеціалізованої компоненти у вигляді конкретизованого потокового графу, який враховує особливості елементної бази;
- розробити структуру спеціалізованої компоненти з максимальною ефективністю використання обладнання, яка враховує всі обмеження та забезпечує обробку даних в реальному масштабі часу;
- розробки інтерфейсу зв'язку між спеціалізованою компонентою та постійною частиною.

Процес проектування спеціалізованих компонентів нейро-орієнтованих комп'ютерних систем реального часу постійно ускладнюється за рахунок підвищення складності виконуваних задач і обмежень, які накладаються в частині габаритів, ваги і споживаної потужності. Крім цього, постійно вимагається зменшення термінів, вартості проектування та підвищення його якості. Забезпечити такі вимоги можливо при інтегрованому підході, який охоплює:

- дослідження та розробку теоретичних основ і методів проектування нейро-орієнтованих комп'ютерних систем реального часу як окремого класу засобів комп'ютерної техніки;
- пошук нових алгоритмічних, структурних, схемотехнічних і конструктивно-технологічних рішень, орієнтованих на НВІС-технології, які

забезпечать задані параметри нейроорієнтованих комп'ютерних систем;

- вибір, а при необхідності проектування, швидкодіючих мало споживаючих НВІС-архітектур, що враховують особливості виконуваних задач і вимог застосування;

- впровадження засобів автоматизованого проектування нейроорієнтованих комп'ютерних систем реального часу, які забезпечать зменшення термінів і підвищать якість проектування.

Структури спеціалізованих компонент повинні в повній мірі використовувати можливості НВІС-технології та забезпечувати обробку інтенсивних потоків інформації в реальному часі [1-3]. Вартість НВІС в основному залежить від площі кристала і кількості виводів. Число зовнішніх виводів НВІС обмежене рівнем технології та розміром кристалу.

Для найповнішого використання переваг НВІС-технології при розробленні спеціалізованих компонент пропонується використовувати такі принципи [3]:

- конвеєризації та просторового паралелізму обробки даних;
- модульності;
- спеціалізації та адаптації апаратно-програмних засобів до структури алгоритмів обробки та інтенсивності надходження даних;
- однорідності та регулярності структури;
- програмованості архітектури шляхом використання репрограмованих ПЛІС.

Використання даних принципів при проектуванні спеціалізованих компонент забезпечить зменшення вартості та термінів їх розроблення. При реалізації спеціалізованої компоненти на базі ПЛІС використовується один з програмних пакетів апаратного проектування (наприклад, Altera Quartus II, або Xilinx ISE). У них, на основі наявних бібліотек, що містять суматори, пристрої множення, блоки обчислення скалярного добутку, реалізації функцій активації, нейрони тощо, проектується спеціалізована компонента на основі схемних редакторів та однієї з мов описання апаратних засобів (наприклад, VHDL).

Після цього здійснюється перевірка коректності роботи розробленої спеціалізованої компоненти за допомогою функціонального та часового моделювання.

Після компіляції розробленого проекту, коли вже відомі необхідні апаратні затрати для реалізації, враховуючи швидкодію і вартість кінцевого пристрою, потрібно вибрати конкретний чіп для реалізації спеціалізованої компоненти.

### **Висновки:**

1. Структура нейро-орієнтованих комп'ютерних систем реального часу складається із двох частин: постійної - обчислювального ядра на базі нейропроцесорів та змінної - спеціалізованих компонентів, кількість яких

залежить від вимог конкретного застосування.

2. Нейро-орієнтовані комп'ютерні системи реального часу доцільно реалізовувати на основі інтегрованого підходу, який охоплює сучасну елементу базу, нові алгоритмічні, структурні та схемотехнічні рішення.

3. Проектування спеціалізованих апаратних компонентів здійснено на таких принципах: конвеєризації та просторового паралелізму обробки даних, однорідності та регулярності структури, спеціалізації та адаптації апаратно-програмних засобів до структури алгоритмів обробки та інтенсивності надходження даних, програмованості архітектури шляхом використання репрограмованих ПЛІС.

4. Розроблено метод синтезу спеціалізованих компонентів нейро-орієнтованої комп'ютерної систем, який враховує особливості елементної бази, вимоги конкретного застосування та забезпечує реалізацію нейроалгоритмів у реальному часі при мінімальних апаратних затратах.

1. Нейроподібні методи, алгоритми та структури обробки сигналів і зображень у реальному часі: монографія / Ю.М. Рашкевич, Р.О. Ткаченко, І.Г. Цмоць, Д.Д. Пелешко. – Львів: Видавництво Львівської політехніки, 2014. -256 с.
2. Інтелектуальні компоненти інтегрованих автоматизованих систем управління: монографія /Медиковський М.О., Ткаченко Р.О., І.Г. Цмоць, Цимбал Ю.В., Дорошенко А.В., Скорохода О.В. – Львів: Видавництво Львівської політехніки, 2015. -280 с.
3. Цмоць І.Г. Інформаційні технології та спеціалізовані засоби обробки сигналів і зображень у реальному часі. – Львів: УАД, 2005.- 227с.
4. С. Хайкин. Нейронные сети: полный курс, 2 –е изд.:Пер. с англ.. –М.:”Вильямс”, 2006. – 1104 с.
5. Галушкин А.И. Нейрокомпьютеры. Кн.3.-М; ИПРЖР,2000.-528с.
6. Проблемно–ориентированные высокопроизводительные вычислительные системы: В.Ф. Гузик, В.Е. Золотовский: Учебное пособие. Таганрог:Изд-во ТРТУ, 1998. 236 с.
7. Уоссермен Ф. Нейрокомпьютерная техника. – М.: Мир,1992. – 259с.
8. А.В. Палагин, В.Н. Опанасенко. Реконфигурируемые вычислительные системы. – К.: Просвіта, 2006.- 280с.
9. Круглов В.В., Борисов В.В. Искусственные нейронные сети. Теория и практика. – М.: Горячая Линия-Телеком, 2002. – 382 с.
10. Николаев А.Б., Фоминых И.Б. Нейросетевые методы анализа и обработки данных. Учебное пособие. - М.: МАДИ (ГТУ), 2003, - 95с.
11. Грибачев В. П. Элементная база аппаратных реализаций нейронных сетей // Компоненты и технологии. 2006. № 8.

*Поступила 19.10.2015р.*