



УДК 004.414.2;004.942;004.02

Член-кореспондент НАН України **В. В. Грицик, І. Г. Цмоць,
В. М. Теслюк**

Методологія системного проектування нейрокомп'ютерних засобів мобільних робототехнічних систем

Розроблено методологію системного проектування нейрокомп'ютерних засобів мобільних робототехнічних систем, яка ґрунтується на інтегральному підході та розроблених принципах побудови таких систем, методах їх проектування та програмно-апаратної реалізації з використанням штучних нейронних мереж і сучасної елементної бази на основі НВІС-структур, що дає змогу підвищити ефективність архітектурних рішень мобільних робототехнічних систем.

Сучасний етап розвитку мобільних робототехнічних систем орієнтований на широке використання нейрокомп'ютерних засобів (НЗ) для оцінювання даних, які надходять від давачів в умовах завод і неповної інформації. Такі нейрокомп'ютерні засоби розробляють на основі паралельних, розподілених та адаптивних технологій обробки даних у реальному часі, які здатні “вчитися” опрацювати дані, діючи в інформаційному середовищі. Особливістю НЗ, які використовуються в робототехнічних системах, є здатність синтезувати правила та здійснювати їх модифікацію в процесі функціонування. В мобільних робототехнічних системах (МРС) вимагається опрацювання у реальному часі різних за інтенсивністю надходження потоків даних на апаратних засобах, що задовольняють обмеження щодо габаритів, енергоспоживання, вартості та часу розробки. Розроблення таких засобів вимагає широкого використання сучасної елементної бази (напівзамовних і замовних НВІС) та побудови нових методів, алгоритмів і НВІС-структур для опрацювання та розпізнавання зображень на основі штучних нейронних мереж (ШНМ). Висока продуктивність і ефективність використання обладнання у нейрокомп'ютерних засобах досягається розпаралеленням і конвєризацією процесів обробки, апаратним відображенням структури алгоритму розв'язання задачі в архітектуру, яка адаптована до інтенсивності надходження потоків даних. Тому актуальною задачею сьогодення є розробка підходів та методів проектування нейрокомп'ютерних засобів мобільних робототехнічних систем на основі сучасної елементної бази.

© В. В. Грицик, І. Г. Цмоць, В. М. Теслюк, 2013

Аналіз існуючих досліджень у галузі проектування НЗ мобільних робототехнічних систем [1–3] показує, що вони мають такі недоліки: не враховують вимоги конкретних застосувань щодо апаратних затрат, габаритів і споживаної потужності; методи проектування та синтезу не враховують особливості сучасної елементної бази; методи, алгоритми обробки та структури апаратних засобів не орієнтовані на НВІС-реалізацію та на роботу в режимі реального часу і не забезпечується висока ефективність використання апаратних засобів.

З проведеного аналізу випливає, що для забезпечення режиму реального часу та високої ефективності використання апаратних засобів необхідно використовувати спеціалізацію, конвеєризацію, просторовий паралелізм і узгодження інтенсивності надходження даних з обчислювальною здатністю апаратних засобів. Тому метою дослідження є розробка методів синтезу високоефективних нейрокомп'ютерних засобів робототехнічних систем.

Принципи побудови нейрокомп'ютерних НВІС-систем. Аналіз останніх досліджень в галузі розроблення РС дає змогу сформулювати основні принципи побудови архітектур нейрокомп'ютерних НВІС-систем, які повинні повною мірою використовувати можливості НВІС-технології, враховувати вартість площі кристала, а також кількість вхідних і вихідних виводів. Число зовнішніх виводів НВІС обмежене рівнем технології та розміром кристала. Для забезпечення зменшення вартості, термінів проектування і високої ефективності використання обладнання в основу побудови нейрокомп'ютерних НВІС-систем пропонується покласти такі принципи: модульності, який передбачає розробку компонентів нейрокомп'ютерних систем (НС) у вигляді функціонально завершених пристроїв (модулів); узгодженості інтенсивності надходження даних з обчислювальною здатністю нейрокомп'ютерних систем; конвеєризації та просторового паралелізму обробки даних; однорідності та регулярності архітектури нейрокомп'ютерних систем; локалізації та спрощення зв'язків між елементами нейрокомп'ютерних систем; спеціалізації та адаптації апаратно-програмних засобів до структури алгоритмів обробки та інтенсивності надходження даних; програмованості архітектури шляхом використання репрограмованих ПЛІС.

Вибір варіанта реалізації нейрокомп'ютерних засобів. Виходячи з вимог, які висуваються до нейрокомп'ютерних засобів, наведемо можливі варіанти їх реалізації: на основі універсальних і функціонально-орієнтованих мікропроцесорів шляхом розробки спеціалізованого програмного забезпечення; на основі універсального обчислювального ядра, доповненого НВІС-пристроями, які реалізують часомісткі алгоритми функціонування ШНМ; у вигляді спеціалізованої алгоритмічної НВІС-системи, архітектура якої апаратно відображає структуру алгоритму.

Перший варіант є доступним для широкого кола користувачів. Перевагою цього варіанту є можливість застосування раніше розроблених програм, а його недоліками є невисока швидкодія, функціональна і структурна надлишковість апаратних засобів.

Другий варіант передбачає поєднання універсальних і спеціальних засобів. Таке поєднання забезпечує високу ефективність використання обладнання при опрацюванні у реальному часі потоків даних за алгоритмами, які є нерегулярними з великою кількістю логічних операцій. При цьому розробка нейрокомп'ютерних засобів з заданими технічними параметрами зводиться до доповнення обчислювального ядра додатковими НВІС-пристроями.

Третій варіант орієнтований на обробку у реальному часі інтенсивних потоків даних за складними нейроалгоритмами. Така реалізація нейроалгоритмів характеризуються введенням додаткового обладнання і відсутністю проміжних пересилок інформації в процесі обчислення, а також спрощенням функції місцевого керування.

Другий та третій варіанти є найперспективнішими для реалізації нейрокомп'ютерних засобів мобільних робототехнічних систем. Дані варіанти передбачають НВІС-реалізацію як базових операцій нейроалгоритмів, так і нейроалгоритмів в цілому.

Для НВІС-реалізацій нейроалгоритми повинні забезпечувати детерміноване переміщення даних, бути добре структурованими та орієнтованими на реалізацію на множині взаємозв'язаних процесорних елементів (ПЕ). Структура та операції, які виконують ПЕ, залежить від вимог, що висуваються до часу обчислення алгоритму. В більшості випадків ПЕ реалізують базові операції нейроалгоритмів, групове підсумовування, обчислення скалярного добутку та передатної функції. При розробці або виборі нейроалгоритмів для НВІС-реалізацій потрібно одночасно враховувати багато взаємопов'язаних факторів.

В першу чергу необхідно, щоб алгоритми були рекурсивними та локально залежними. В рекурсивному алгоритмі всі ПЕ повинні виконувати приблизно однакові операції. При реалізації рекурсивного алгоритму кожний із ПЕ буде повторювати виконання фіксованого набору операцій над послідовністю даних, що надходять. Ефективність відображення алгоритму на ПЕ безпосередньо пов'язана із способом декомпозиції розв'язання задачі перетворення на незалежні базові операції, що виконуються паралельно, або на залежні, що виконуються у конвеєрному режимі. Рекурсивні алгоритми можна розділити на два класи з локальними (обміни здійснюються тільки між найближчими сусідніми ПЕ) та глобальними пересилками даних.

Особливості проектування нейрокомп'ютерних засобів. Проектування нейрокомп'ютерних засобів складається з таких етапів: вибору та розробки методів і алгоритмів функціонування нейрокомп'ютерних засобів; визначення основних параметрів апаратних засобів; переходу від алгоритму до структури апаратних засобів [3].

Розробку нейрокомп'ютерних засобів мобільних робототехнічних систем доцільно здійснювати на основі компонентно-ієрархічного підходу, який передбачає поділ процесу розробки на ієрархічні рівні та види забезпечення (алгоритмічне, апаратне та програмне). Для реалізації такого підходу використовується метод декомпозиції, який передбачає розбиття засобів на окремі компоненти, з яких синтезуються нейрокомп'ютерні засоби для вимог конкретного застосування. На кожному рівні ієрархії розв'язуються задачі відповідної складності, які характеризуються як одиницями інформації, так і алгоритмами обробки.

За складністю опрацювання даних компоненти нейрокомп'ютерних систем можна розділити на чотири ієрархічні рівні. Збільшенню номера рівня ієрархії відповідає збільшення деталізації алгоритмічних, апаратних і програмних засобів. При цьому на вищих рівнях ієрархії одиниці інформації, алгоритми, програмні та апаратні засоби являють собою впорядковані сукупності одиниць інформації та композиції алгоритмів, програмних і апаратних засобів нижчих рівнів ієрархії (табл. 1).

Методологія послідовної декомпозиції, яка використовується при компонентно-ієрархічному підході до розробки нейрокомп'ютерних засобів мобільних робототехнічних систем, відображає процес проектування "зверху вниз".

На першому ієрархічному рівні розробляються методи, алгоритми опрацювання даних, структури апаратних та програмних засобів. Здійснюється відображення алгоритму роботи у вигляді функціонального графа $F = (\Phi, \Gamma)$, де $\Phi = \{\Phi_1, \Phi_2, \dots, \Phi_n\}$ — множина функціональних операторів; Γ — закон відображення зв'язків між операторами [5].

Другий рівень ієрархії складають методи, алгоритми та структури підсистеми керування рухом і технічного зору, які реалізуються на базі ШНМ, нейропроцесорів, процесорів обробки сигналів, мікроконтролерів та паралельної пам'яті [2].

Третій ієрархічний рівень складається з методів, алгоритмів та структур компонентів, які реалізують базові макрооперації алгоритмів керування рухом і технічного зору та нейро-алгоритмів. Для реалізації елементів третього рівня розробляються паралельні алгоритми, НВІС-структури та програми.

До четвертого рівня ієрархії відносяться елементи, які реалізують операції нейронного елемента, функції активації, радіально-базисної функції, скалярного добутку, групового підсумовування. В функціональному і структурному відношеннях елементи четвертого рівня ґрунтуються на елементарних арифметичних операціях.

Компонентно-ієрархічний підхід до розробки нейрокомп'ютерних засобів мобільних робототехнічних систем можна описати за допомогою такого виразу:

$$C_{\text{MIT}}^1 = \bigcup_{i=1}^n C_{\text{MIT}}^{2i} \bigcup_{j=1}^m C_{\text{MIT}}^{3j} \bigcup_{p=1}^h C_{\text{MIT}}^{4p},$$

де C_{MIT}^{2i} , C_{MIT}^{3j} , C_{MIT}^{4p} — засоби відповідно другого, третього і четвертого ієрархічних рівнів; n — кількість типів підсистем; m — кількість типів компонентів; h — кількість типів операційних пристроїв.

Розробка архітектури нейрокомп'ютерних НВІС-систем. Пропонується розробку архітектури нейрокомп'ютерних НВІС-систем здійснювати на основі інтегрованого підходу, який ґрунтується на можливостях сучасної елементної бази та охоплює структури, методи, алгоритми функціонування компонентів і нейрокомп'ютерних систем, враховує вимоги конкретних застосувань та інтенсивності надходження даних.

При цьому задача проектування компонентів і нейрокомп'ютерних систем зводиться до формування множин вимог $\mathbf{R} = \{R_1, R_2, \dots, R_k\}$, характеристик $\mathbf{H} = \{H_1, H_2, \dots, H_m\}$ і обмежень $\mathbf{B} = \{B_1, B_2, \dots, B_k\}$ та знаходження такого вектора $\mathbf{H}^* = [H_1^*, H_2^*, \dots, H_m^*]$,

Таблиця 1. Рівні та види розробок нейрокомп'ютерних засобів

Ієрархічний рівень	Види забезпечення та виконувані розробки		
	Алгоритмічне	Апаратне	Програмне
1-й	Методи та алгоритми функціонування НС	Структура нейрокомп'ютерної системи	Структура програмних засобів НС
2-й	Методи та алгоритми функціонування підсистеми керування рухом та підсистеми технічного зору	Структури ШНМ, нейропроцесорів, процесорів обробки сигналів, мікроконтролерів, паралельної пам'яті	Програми реалізації ШНМ, програм керування рухом, оцінювання даних, компресування та розпізнавання зображень
3-й	Методи та алгоритми реалізації компонентів підсистем керування рухом та технічного зору	Структури базових компонентів підсистем керування рухом та технічного зору нейронних елементів, пристроїв для реалізації макрооперацій нейроалгоритмів	Програми реалізацій шарів ШНМ, налаштування ШНМ, обміну, опитування давачів, макрооперацій нейроалгоритмів
4-й	Методи та алгоритми реалізації нейронного елемента, функції активації, радіально-базисної функції, скалярного добутку, групового підсумовування	Структури нейронного елемента, операційних пристроїв для реалізації радіально-базисної функції, скалярного добутку, групового підсумовування	Програми реалізацій нейронного елемента, радіально-базисної функції, скалярного добутку, групового підсумовування

$H_i^* = f_i(R, H, B)$, $i = 1, \dots, m$, який забезпечить максимальне значення ефективності використання обладнання $E = \max f(R, H^*, B)$.

Множину вимог \mathbf{R} складає: R_1 — кількість каналів надходження даних m_d ; R_2 — розрядність каналів надходження даних n_d ; R_3 — частота надходження даних F_d ; R_4 — швидкість елементної бази, яка визначається часом затримки вентиля t_b ; R_5 — кількість елементів (слів) вхідного масиву N ; R_6 — розрядність вхідного слова n . Множину характеристик \mathbf{H} становлять: H_1 — загальна кількість зв'язків Z ; H_2 — просторова зв'язкова віддаль Δj ; H_3 — конвеєрний такт t_k ; H_4 — витрати обладнання W ; H_5 — кількість типів функціональних вузлів s ; H_6 — кількість каналів введення $m_{\text{вв}}$; H_7 — розрядність каналів введення $n_{\text{вв}}$; H_8 — кількість виводів інтерфейсу зв'язку Y . Обмеження \mathbf{B} , які необхідно враховувати при проектуванні НВІС-систем реального часу, такі: B_1 — точність обчислення, що визначається розрядністю результату n_p ; B_2 — час обчислення $T_{\text{обч}}$, повинен бути $T_{\text{обч}} \leq T_{\text{обм}}$, де $T_{\text{обч}} = t_k N n / (m_{\text{вв}} n_{\text{вв}})$; $T_{\text{обм}}$ — час обміну, який визначається так: $T_{\text{обм}} = N n / (F_d m_d n_d)$.

Для вибору варіанта структури нейрокомп'ютерної НВІС-системи застосовується критерій ефективності використання обладнання E , який враховує кількість виводів інтерфейсу, однорідність структури, кількість і локальність зв'язків, зв'язує продуктивність з витратами обладнання та дає оцінку елементам (вентиліям) компонента за продуктивністю [2]. Кількісна величина ефективності використання обладнання для такого апаратного засобу визначається таким чином:

$$E = \frac{G m_k n_k}{t_k N n \left(k_1 \sum_{i=1}^s W_{\text{ФУ}_i} d_i + k_2 Q + k_3 Y \right)},$$

де G — складність алгоритму опрацювання; t_k — конвеєрний такт; $W_{\text{ФУ}_i}$ — витрати обладнання у вентилях на реалізацію i -го функціонального вузла; d_i — кількість функціональних вузлів i -го типу; k_1 — коефіцієнт врахування однорідності $k_1 = f(s)$; k_2 — коефіцієнт врахування регулярності зв'язків $k_2 = f(\Delta j)$; k_3 — коефіцієнт врахування кількості виводів інтерфейсу зв'язку $k_3 = f(Y)$.

Конвеєрний такт t_k визначається за формулою $t_k = \sum_j^l \max t_b$, де l — кількість послідовно з'єднаних вентилів у найповільнішій сходинці конвеєра, а Δj — як різниця просторових індексів.

До основних параметрів оцінки апаратних засобів реального часу, крім витрат обладнання, швидкодії, ефективності використання обладнання, пропонується використовувати обчислювальну здатність. Для обробки потоків даних у реальному часі доцільно застосовувати синхронні структури з конвеєрною реалізацією графів нейроалгоритмів, в яких здійснюється суміщення у часі виконання функціональних операторів нейроалгоритму над різними даними. Конвеєризація нейрокомп'ютерної системи передбачає розділення її на сходинки шляхом введення буферної пам'яті. При цьому кожна сходинка конвеєра складається з двох компонентів: буферної пам'яті та операційних пристроїв, що реалізують функціональні оператори ярусу. Для забезпечення високої швидкодії та ефективності використання обладнання функціональні оператори, які реалізуються у сходинках конвеєра, мають бути простими та мати приблизно однаковий час реалізації.

Перехід від алгоритму розв'язання задачі до структури нейрокомп'ютерної НВІС-системи, яка працює в реальному часі, формально зводиться до мінімізації витрат обладнання

$W_{\text{НС}} = W_{\text{ІП}} + W_{\text{П}} + W_{\text{ПУ}} + \sum_{i=1}^k W_{\text{ПЕ}_i} m_i$, де $W_{\text{НС}}$, $W_{\text{ПУ}}$, $W_{\text{ІП}}$, $W_{\text{П}}$, $W_{\text{ПЕ}}$ — витрати обладнання відповідно на нейрокомп'ютерну систему, пристрої керування, інтерфейсні пристрої, пам'ять; k — кількість типів ПЕ, i -й ПЕ; m_i — кількість ПЕ i -го типу, при забезпеченні такої умови: $Nn/(F_d m_d n_d) \geq t_k Nn/(m_{\text{ВВ}} n_{\text{ВВ}})$.

Основними шляхами мінімізації апаратних затрат на реалізацію нейрокомп'ютерної НВІС-системи є:

вибір ефективних методів і алгоритмів реалізації функціональних операторів нейроалгоритмів;

зменшення розрядності операційних пристроїв, пам'яті, кількості і розрядності каналів передачі даних;

узгодження інтенсивності надходження даних із обчислювальною здатністю нейрокомп'ютерної НВІС-системи.

На закінчення зробимо такі висновки.

1. Розробку нейрокомп'ютерних засобів мобільних робототехнічних систем доцільно здійснювати на основі інтегрованого підходу, який охоплює сучасну елементну базу, методи та алгоритми реалізації нейроалгоритмів, архітектури компонентів і систем, враховує вимоги конкретних застосувань, інтенсивності надходження даних і ґрунтується на таких принципах побудови: конвеєризації та просторового паралелізму обробки даних; модульності; однорідності та регулярності архітектури; спеціалізації та адаптації апаратно-програмних засобів до структури алгоритмів обробки та інтенсивності надходження даних.

2. Для вибору архітектури нейрокомп'ютерних НВІС-систем запропоновано застосовувати критерій ефективності використання обладнання, який враховує кількість виводів інтерфейсу, однорідність структури, кількість і локальність зв'язків, зв'яже продуктивність з витратами обладнання та дає оцінку елементам системи за продуктивністю.

3. Основними шляхами підвищення ефективності використання обладнання нейрокомп'ютерних НВІС-систем є: вибір ефективних методів і алгоритмів реалізації для НВІС-реалізації; зменшення розрядності операційних пристроїв, пам'яті, кількості і розрядності каналів передачі даних; узгодження інтенсивності надходження даних із обчислювальною здатністю апаратних засобів на всіх рівнях.

4. Визначено, що узгодження інтенсивності надходження даних із обчислювальною здатністю НВІС-систем можна здійснювати так: зміною тривалості конвеєрного такту, кількості і розрядності каналів надходження даних в операційних пристроях.

5. Основними етапами синтезу нейрокомп'ютерних НВІС-систем є: вибір та розробка методів і алгоритмів узгоджено-паралельної обробки; визначення основних параметрів апаратних засобів; перехід від алгоритму до узгодженої паралельної структури.

1. Резник А. М., Куссуль М. Э., Сычов А. С. и др. Система автоматизированного проектирования модульных нейронных сетей САД МNN // Мат. машины и системы. – 2002. – № 3. – С. 28–36.
2. Подураев Ю. В., Кулешов В. С. Принципы построения и современные тенденции развития мехатронных систем // Мехатроника. – 2000. – № 1.
3. Круглов В. В., Борисов В. В. Искусственные нейронные сети. Теория и практика. – 2-е изд., стереотип. – Москва: Горячая линия-Телеком, 2002. – 382 с.
4. Хайкин С. Нейронные сети: полный курс. – 2-е изд. / Пер. с англ. – Москва: Вильямс, 2006. – 1104 с.
5. Галушкин А. И. Нейрокомпьютеры. Кн. 3. – Москва: ИПРЖР, 2000. – 528 с.
6. Осовский С. Нейронные сети для обработки информации / Пер. с польского. – Москва: Финансы и статистика, 2002. – 344 с.

7. Палагин А. В., Опанасенко В. Н. Реконфигурируемые вычислительные системы. – Киев: Просвіта, 2006. – 280 с.
8. Ткаченко Р. О. Новая парадигма штучних нейронних мереж прямого поширення // Вісн. Держ. ун-ту “Львівська політехніка”: Комп’ютерна інженерія та інформаційні технології. – 1999. – № 386. – С. 43–54.
9. Цмоць І. Г. Інформаційні технології та спеціалізовані засоби обробки сигналів і зображень у реальному часі. – Львів: УАД, 2005. – 227 с.
10. Грицик В. В., Ткаченко Р. О. Нові підходи до навчання штучних нейромереж // Доп. НАН України. – 2002. – № 11. – С. 59–65.
11. Грушицкий Р. И., Мурсаев А. Х., Угрюмов Е. П. Проектирование систем на микросхемах программируемой логики. – Ст.-Петербург: БХВ-Петербург, 2002. – 608 с.
12. Воеводин В. В., Воеводин Вл. В. Параллельные вычисления. – Ст.-Петербург: БХВ-Петербург, 2002. – 608 с.
13. Немнюгин С. А., Стесик О. Л. Параллельное программирование для многопроцессорных систем. – Ст.-Петербург: БХВ-Петербург, 2002. – 400 с.
14. Касьянов В. Н., Евстигнеев В. А. Графы в программировании: обработка, визуализация и применение. – Ст.-Петербург: БХВ – Петербург, 2003. – 1104 с.
15. Ткаченко Р. О., Ткаченко П. Р., Цмоць І. Г. Апаратна реалізація багат шарових перцептронів з неінтеграційним навчанням: Зб. наук. праць. – Київ: Ін-т проблем моделювання в енергетиці НАН України. – 2005. – Вип. 29. – С. 103–113.

Національний університет “Львівська політехніка”

Надійшло до редакції 10.05.2012

Член-корреспондент НАН України **В. В. Грицик, І. Г. Цмоць, В. Н. Теслюк**

Методология системного проектирования нейрокомпьютерных средств мобильных робототехнических систем

Разработана методология системного проектирования нейрокомпьютерных средств мобильных робототехнических систем, которая базируется на интегральном подходе и разработанных принципах построения таких систем, методах их проектирования и программно-аппаратной реализации с использованием искусственных нейронных сетей и современной элементной базы на основании СВИС-структур, что дает возможность повысить эффективность архитектурных решений мобильных робототехнических систем.

Corresponding Member of the NAS of Ukraine **V. V. Grytsyk, I. G. Tsmots, V. M. Teslyuk**

Methodology of the system design of neuro-computer instruments for mobile robotic systems

A methodology of the system design of neuro-computer instruments for mobile robotic systems is developed. It is based on the integral approach and principles of the design of such systems, methods of their design and their program-practical realization with the use of artificial neuron systems, and modern elements based upon HBIC-structures that enables to increase the efficiency of architectural decisions for mobile robotic systems.