

МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ТА ОБЧИСЛЮВАЛЬНІ МЕТОДИ

УДК 004.272

Т. Б. МАРТИНЮК, А. В. КОЖЕМ'ЯКО, Н. О. ДЕНИСЮК, Т. Ю. ПОЗДНЯКОВА

Вінницький національний технічний університет, м. Вінниця

АНАЛІЗ ОПЕРАЦІЙНОГО БАЗИСУ ДЛЯ НЕЙРОМЕРЕЖЕВИХ ІНТЕЛЕКТУАЛЬНИХ СИСТЕМ

Анотація. Розглянуто особливості реалізації комп'ютерних засобів у складі інтелектуальних систем різного призначення, зокрема за нейромережевою технологією. Обґрунтовано вибір базових обчислювальних процедур та операцій для інтелектуальних систем. Показано, що широко розповсюдженою операцією в задачах обробки, аналізу сигналів і зображень та розпізнавання образів є векторно-матричне перемноження. Ця операція реалізується шляхом формування парних добутків та багатооперандного підсумовування. Наведено перспективні концепції реалізації операційного базису для інтелектуальних систем, з використанням різницево-зрізової обробки векторних масивів даних.

Ключові слова: інтелектуальні системи, операційний базис, обчислювальні процедури та операції.

Аннотация. Рассмотрены особенности реализации компьютерных средств в составе интеллектуальных систем различного назначения, в частности по нейросетевой технологии. Обоснован выбор базовых вычислительных процедур и операций для интеллектуальных систем. Показано, что широко распространенной операцией в задачах обработки, анализа сигналов и изображений и распознавания образов является векторно-матричное умножение. Эта операция реализуется путем формирования парных произведений и многооперандного суммирования. Приведены перспективные концепции реализации операционного базиса для интеллектуальных систем, с использованием разностно-срезовой обработки векторных массивов данных.

Ключевые слова: интеллектуальные системы, операционный базис, вычислительные процедуры и операции.

Abstract. It is dealt with computer assets of intelligent systems for various purpose, in particular with neural network technology. It is spoken in detail basic calculation procedures and operations for intelligent systems. In problems of processing, analysis of signals and images and pattern recognition the most common operation is vector-matrix multiplication. It is realization by formation pair of products and many operand summing. The promising concept operational basis for the implementation of intelligent systems is shown.

Key words: intelligent systems, operational basis, processing procedures and operations

Вступ

Концептуальною базою для інтелектуалізації комп'ютерних систем є методи штучного інтелекту з використанням нейроінформаційних технологій [1, 2]. З іншого боку в роботі [1] визначено, що штучні нейромережі (ШНМ), в свою чергу, представляють собою окремий клас інтелектуальних систем. Причому, саме універсальні апроксимуючі властивості ШНМ можуть бути використані при створенні нового покоління інтелектуальних комп'ютерних систем [3].

Актуальність

З публікацій видно, що потреба в інтелектуальних системах різного призначення не зменшується [1-8]. Про це свідчать розробки в області створення інтелектуальних телекомунікаційних систем, інтелектуальних систем керування мобільними роботами, засобів захисту інформації в інтелектуальних системах, моделювання "нервової діяльності" в інтелектуальних системах, інтелектуальних систем ідентифікації, геоінформаційної інтелектуальної системи моніторингу стану сільськогосподарських культур, системи інтелектуального керування трафіком у комп'ютерних мережах тощо [2, 5, 6, 8]. При цьому обґрунтування вибору конкретної нейротехнології в процесі розроблення інтелектуальних засобів визначається класом розв'язуваних задач [1].

Мета

Метою даної роботи є аналіз операційного базису нейромережевих інтелектуальних систем для обґрунтування вибору функціонально-повного набору обчислювальних операцій на нових принципах їх реалізації.

Постановка задачі

Відомо, що ШНМ представляють собою "величезний розподілений паралельний процесор, що складається з елементарних одиниць оброблення інформації, які накопичують експериментальні знання і передають їх для подальшого оброблення" [1,2]. І тільки властивості ШНМ, пов'язані з їх здатністю до навчання та узагальнення отриманих знань, наділяють їх рисами штучного інтелекту [1].

Отже, зважаючи на різноплановість застосування нейротехнологій, зокрема ШНМ, у вирішенні практичних задач, правомірним є використання такого позначення, як "обчислювальні засоби інтелектуальних систем", орієнтуючись саме на застосування нейроструктур. На таких же засадах застосовують поняття "обчислювальний інтелект" [7].

Серед найбільш актуальних задач, що розв'язуються інтелектуальними системами, варто відмітити такі [1,2,5,6,8]: ефективне керування дистанційними об'єктами (наприклад, керування мобільними роботами); реалізація "технічного мозку" для навчених систем керування промисловими роботами; природньо наближене моделювання функцій нервової системи при біомедичних дослідженнях; медичне експрес-діагностування; розпізнавання зорових та мовленнєвих образів;

інтелектуалізація систем ідентифікації; підвищення ефективності систем захисту інформації в комп'ютерних та телекомунікаційних мережах.

При розробці методів оброблення та аналізу даних для обчислювальних засобів інтелектуальних систем необхідно враховувати такі особливості та обмеження, що пов'язані з характером наведених вище розв'язуваних задач: задачі мають яскраво виражений природний паралелізм подання та оброблення даних; апіорна інформація про характеристики розподілу великих масивів цифрових даних відсутня або містить приблизні значення; оброблення та аналіз даних і розпізнавання образів повинні виконуватися в реальному часі, що обумовлено необхідністю оперативного прийняття рішень; пріоритетною повинна бути апаратна реалізація засобів через компактність і мобільність їх застосування.

Отже, основним напрямком покращення алгоритмічних показників багатьох процесів оброблення та аналізу сигналів і зображень, а також розпізнавання образів є збільшення рівня паралелізму при обробленні масивів даних і, зокрема векторних масивів даних, в процесі виконання базових операцій множення і багатооперандного підсумовування.

Аналіз базових операцій обчислювальних систем

Аналіз сучасних комп'ютерних систем показав, що вони реалізують у сукупності всі базові інформаційні технології, зокрема оброблення та аналіз сигналів і зображень, а також розпізнавання образів. З іншого боку, аналіз складових найбільш розповсюджених інформаційних технологій [1, 2, 9, 10] дозволяє навести класифікаційну схему базових обчислювальних процедур та операцій (рис. 1) і визначити такі базові процедури, як фільтрація, згортка, кореляція, класифікація, паралельний пошук і нейрообчислення. Ці процедури потребують виконання таких базових операцій, як векторно-матричне перемноження, сортування, визначення екстремальних значень, ранжування, порівняння з еталоном, паралельне порівняння.

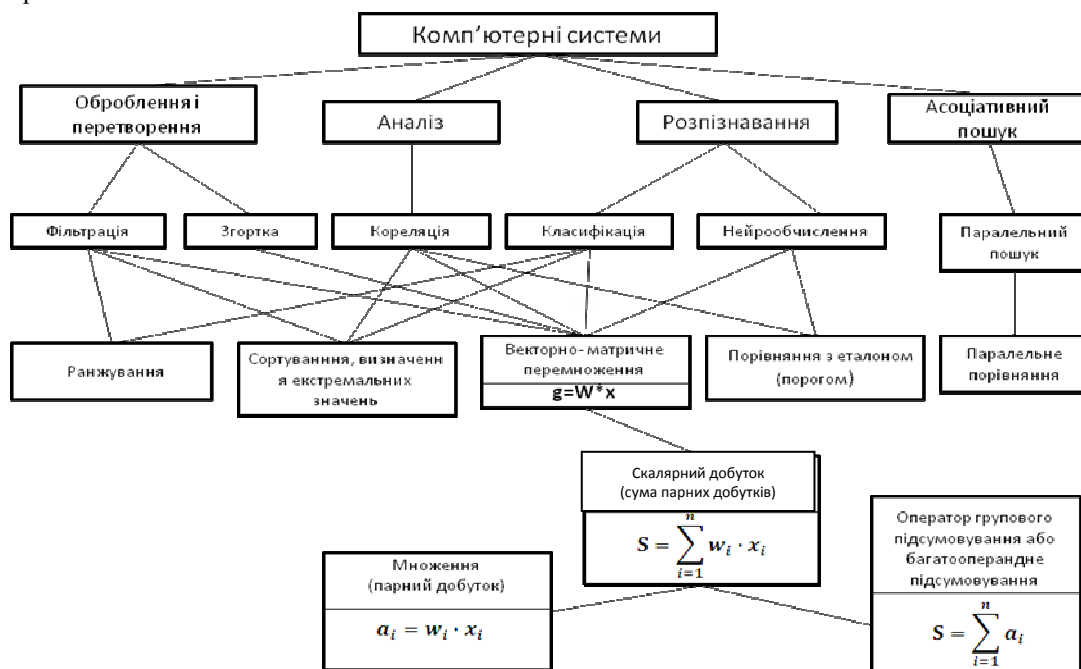


Рисунок 1 – Класифікаційна схема базових обчислювальних процедур та операцій

Разом з тим, серед особливостей нейрообчислень слід відзначити [10]:

- великий обсяг обчислень з переважанням обчислювальних операцій над логічними;
- доцільність використання векторного оброблення даних;
- використання двох видів паралелізму – просторового і часового;
- доцільність розпаралелювання та конвеєризації процесу оброблення;
- базові операції типу множення-підсумовування.

Отже, наочним є факт, що найбільшу складову серед нейрообчислень має формування добутку вектора біжучого стану мережі на матрицю міжнейронних зв'язків. Тому для прискорення цієї нейрооперації застосовують апаратні нейроакселератори [1, 2, 10], наприклад, багатопроекторний варіант потокового нейрокомп'ютера на базі трансп'ютерів IMS T805.

Все це дозволяє серед наведених вище базових операцій визначити особливе місце операції множення матриці на вектор, яка широко використовується не тільки при моделюванні НМ, але й при

розв'язанні прикладних задач у системах оброблення й аналізу зображень, при розпізнаванні образів, у машинній графіці, для реалізації обчислень з плаваючою комою.

У свою чергу, базову операцію векторно-матричного перемноження вигляду

$$\mathbf{g} = \mathbf{W} \cdot \mathbf{x}, \quad (1)$$

можна розкласти на такі складові, як скалярний добуток або суму парних добутоків вигляду

$$S = \sum_{i=1}^n w_i \cdot x_i, \quad (2)$$

а в подальшому відповідно на операцію множення (парний добуток) вигляду

$$a_i = w_i \cdot x_i \quad (3)$$

та оператор групового підсумовування (ОГП) вигляду

$$S = \sum_{i=1}^n a_i \quad (4)$$

Про актуальність реалізації операції (1) свідчать наведені у роботі [11] результати паралельно-конвеєрної реалізації задачі множення матриці на потік векторів на реконфігурованих обчислювальних системах (РОС), тобто на ПЛІС. Так, експериментальні дослідження показали, що РОС на базі 16 ПЛІС XC4VLX40 з тактовою частотою 160 МГц мають продуктивність 21,7 ГФлопс, а на базі 16 ПЛІС XC4VLX80 з тактовою частотою 250 МГц - 70,9 ГФлопс, що у порівнянні з продуктивністю ПК Intel(R) Core(TM) 2 CPU 4400@2.00GHz складає вигравш у часі відповідно у 54 рази і 176 разів.

У свою чергу, паралельні методи, алгоритми та НВІС-структури обчислення сум парних добутоків (2) достатньо докладно розглянуто у роботі [10], де наведено аналітичні вирази для оцінювання основних характеристик пристроїв обчислення сум парних добутоків з врахуванням кількості операндів та їх розрядності.

Разом з тим не менш актуальними є задачі ефективної реалізації операції множення (3) та ОГП (4). У роботах [12, 13] наведено переваги апаратної реалізації операції множення, зокрема на ПЛІС. У статті [13] наведено оцінки складності в базисі ПЛІС FPGA серії XC4000E для типів помножувачів з послідовною, паралельно-послідовною та табличною архітектурами. Перевагу мають помножувачі з паралельно-послідовною архітектурою для чисел великої розмірності ($n = 16; 32$). У роботі [12] обґрунтовано вибір для реалізації на ПЛІС схеми помножувача на базі алгоритму Бута і доведено, що найбільш придатним в цьому випадку для згортки часткових добутоків є ієрархічне "дерево" багаторозрядних масштабуючих суматорів.

У порівнянні з множенням операція багатооперандного підсумовування (4) має значно менше можливостей для підвищення рівня паралелізму [14]. У роботі [14] детально розглянуто і досліджено розпаралелювання реалізації ОГП (4). Найбільш прийнятним способом є почисловий паралельно-послідовний, відомий як алгоритм логарифмічного підсумовування [9, 15]. Він зорієнтований на використання відомої структури суматорів у вигляді "дерева", що дозволяє організувати виконання будь-якої операції типу "згортки" за рахунок її максимального розпаралелювання. В даному випадку довжина критичного шляху у відповідних інформаційних графах зводиться до величини порядку $t \cdot \log_2 n$, де n – розмірність масиву даних, t – час виконання елементарної операції [9]. Для цього способу характерним є виконання сукупності бінарних операцій підсумовування, що відрізняє його від порозрядного паралельно-послідовного способу, який використовує обчислення порозрядних сум (розрядних зрізів) сукупності доданків [15]. Останній відомий як спосіб багатооперандного підсумовування, в основі якого лежить перетворення багаторядного коду в однорядний [16].

При апаратній реалізації швидкодіючих багатовхідних суматорів можливий вибір або комбінація двох варіантів: на базі цифрового сигнального процесора (ЦСП) або на ПЛІС. При використанні ЦСП DSP48E маємо тактову частоту 400 МГц і логічну ємність – 1 блок, а при використанні ПЛІС Xilinx Virtex 5 маємо тактову частоту 154 МГц і логічну ємність – 44 комбінаційних логічних блоків (КЛБ, LUT) [17]. В даному випадку вигравш у швидкодії досягається за рахунок ефективного трасування сигнальних ліній в рамках ЦСП.

Крім того, особливу увагу варто приділити саме багатооперандному підсумовуванню, оскільки доведено, що пара операцій $(\min, +)$ утворює напівкільце на розширеній множині дійсних чисел (включно із " ∞ ") [18]. Пара операцій $(\min, +)$ задовольняє умовам, що визначають напівкільце, а саме [18]:

- 1) $\min(x, y) = \min(y, x)$;
- 2) $\min(\min(x, y), z) = \min(x, \min(y, z))$;
- 3) $x + y = y + x$;
- 4) $(x + y) + z = x + (y + z)$;
- 5) $x + \min(y, z) = \min(x + y, x + z)$;
- 6) множина $\mathbb{R} \cup \{\infty\}$ містить " ∞ ", для якого при будь-якому x виконуються рівності $\min(x, \infty) = x$, $x + \infty = x$;
- 7) множина $\mathbb{R} \cup \{\infty\}$ містить 0, яке при будь-якому x задовольняє рівність $x + 0 = x$.

Таким чином, подальше дослідження з метою підвищення ефективності виконання базових процедур для інтелектуальних систем базується на розробленні та вдосконаленні базових обчислювальних операцій на нових принципах організації такої операції, як багатооперандне підсумовування. З цією метою серед перспективних концепцій розвитку інформаційних технологій оброблення і аналізу даних та розпізнавання образів в інтелектуальних системах необхідно виділити застосування принципів око-процесорного оброблення цифрових даних, що дозволить [3, 19]: досягти глибокого розпаралелювання при обробленні одно- і двовимірних масивів даних; сумістити виконання базових нейрооперацій (підсумовування зважених вхідних сигналів і порогове порівняння); розширити функціональні можливості оброблення елементів одно- і двовимірних масивів даних, наприклад, за рахунок їх сортування та ранжування, сумішених зі згорткою елементів масиву.

Висновки

1. На основі здійсненого аналізу операційного базису для інтелектуальних систем відзначено, що найбільш розповсюдженою операцією в задачах оброблення, аналізу сигналів та зображень і розпізнавання образів є операція векторно-матричного перемноження, яка, в свою чергу, реалізується в процесі формування парних добутків та їх групового підсумовування (багатооперандного підсумовування)
2. Шляхом обґрунтування базових операцій для комп'ютерних засобів інтелектуальних систем показано обмеженість в алгоритмічно-структурному аспекті класичного варіанта розпаралелювання операції групового підсумовування.
3. При проведенні аналізу визначено актуальну задачу розроблення такого методу багатооперандного підсумовування, який би забезпечив не тільки високий рівень паралелізму, але й можливість суміщення при виконанні декількох операцій та отримання проміжних результатів.

Список літератури

1. Осовский С. Нейронные сети для обработки информации / С. Осовский; пер. с польск. И.Д. Рудинского. – М.: Финансы и статистика, 2004. – 344 с. – ISBN 5-279-02567-4.
2. Комашинский В.И. Нейронные сети и их применение в системах управления и связи / В.И. Комашинский, Д.А. Смирнов. – М.: Горячая линия-Телеком, 2003. – 94 с. – ISBN 5-93517-094-9.
3. Кожем'яко В.П. Наукова концепція образного відео-комп'ютера око-процесорного типу в контексті сучасної методології штучного інтелекту / В.П. Кожем'яко, А.А. Яровий // Оптико-електронні інформаційно-енергетичні технології. – 2001. – № 2. – С. 84-89. – ISSN 1681-7893.
4. Палагин А.В. Системная интеграция средств компьютерной техники / А.В. Палагин, Ю.С. Яковлев. – Вінниця: УНІВЕРСУМ-Вінниця, 2005. – 680 с. – ISBN 966-641-140-7.
5. Гаврилов А.В. Архитектура гибридной интеллектуальной системы управления мобильного робота / А.В. Гаврилов, В.В. Губарев, К. Х. Джо, Х. Х. Ли // Вестник Новосибирского ГТУ. – 2004. – № 2. – С. 3-13.
6. Гильгурт С.Я. Программно-аппаратная защита данных в распределённых интеллектуальных системах / С. Я. Гильгурт, А.К. Гиранова // Искусственный интеллект. – 2010. – № 3. – С. 706-711. – ISSN 1561-5359.
7. Rutkowski L. Computational Intelligence. Methods and Techniques. Berlin-Heidelberg: Springer-Verlag, 2008, 514 p.
8. Буков А.А. Технические нервные системы. Обучаемые системы управления со зрением для промышленных роботов / А.А. Буков. – Липецк: Изд-во Липецк. гос. техн. ун-та, 2001. – 223 с.
9. Прэтт У. Цифровая обработка изображений / У. Прэтт; пер. с англ. – В 2-х кн. – М.: Мир, 1982. – 792 с.
10. Цмоць І.Г. Інформаційні технології та спеціалізовані засоби обробки сигналів і зображень у реальному часі / І.Г. Цмоць. – Львів: Видавництво УАД, 2005. – 228 с. – ISBN 966-322-024-4.

11. Трунов Г.Л. Параллельно-конвейерная реализация задач и умножения матрицы на поток векторов на реконфигурируемых вычислительных системах / Г.Л. Трунов, А.Г. Коваленко // Искусственный интеллект. – 2008. – № 3. – С.742-749. – ISSN 1561-5359.
12. Стешенко В.Б. ПЛИС фирмы Altera: элементная база, система проектирования и языки описания аппаратуры / В.Б. Стешенко. – М.: Издательский дом "Додэка-XXI", 2002. – 576 с. – ISBN 5-94120-033-1.
13. Особенности реализации операции умножения на ПЛИС / МоЧжо Чо // Современные наукоемкие технологии. – 2008. – № 4 [Электронный ресурс] – Режим доступа: http://www.rae.ru/fs/?section=content&op=show_article&article_id=7791584.
14. Справочник по цифровой вычислительной технике; под ред. В.Н. Малиновского. – К.: Техника, 1980. – 320 с.
15. Мартинюк Т.Б. Рекурсивні алгоритми багатоперандної обробки інформації / Т.Б. Мартинюк. – Вінниця: УНІВЕРСУМ-Вінниця, 2000. – 216 с. – ISBN 966-7199-98-3.
16. Гамаюн В.П. Способ ускоренного преобразования многорядного кода в однорядный / В.П. Гамаюн // Управляющие системы и машины. – 1995. – № 4/5. – С. 10-14. – ISSN 0130-5395.
17. Зианбетов Э.И. Исследование специализированных структур цифровой обработки сигналов в составе ПЛИС серии Virtex / Э.И. Зианбетов, В.П. Малахов, В.С. Ситников // Современные информационные и электронные технологии: Девятая междунар. науч.-практ. конф., 19-23 мая 2008 г.: труды. – Одесса, 2008. – С. 133.
18. Шлезингер М. Десять лекций по статистическому и структурному распознаванию / М. Шлезингер, В. Главач. – К.: Наук. думка, 2004. – 536 с. – ISBN 966-00-0341-2.
19. Кожем'яко В.П. Квантові перетворювачі на оптоелектронних логіко-часових середовищах для око-процесорної обробки зображень / В.П. Кожем'яко, Т.Б. Мартинюк, О.І. Суприган, Д.І. Клімкіна. – Вінниця: УНІВЕРСУМ-Вінниця, 2007. – 126 с. – ISBN 978-966-641-219-8.

Відомості про авторів

Мартинюк Тетяна Борисівна – д.т.н., доцент, професор кафедри лазерної та оптикоелектронної техніки, Вінницький національний технічний університет.

Кожем'яко Андрій Вікторович - к.т.н., доцент, доцент кафедри лазерної та оптикоелектронної техніки, Вінницький національний технічний університет.

Денисюк Наталія Олексіївна – здобувач кафедри лазерної та оптикоелектронної техніки, Вінницький національний технічний університет.

Позднякова Тетяна Юріївна - студентка магістратури кафедри лазерної та оптикоелектронної техніки, Вінницький національний технічний університет.