

АВТОМАТИЗАЦІЯ ТА ІНТЕЛЕКТУАЛІЗАЦІЯ ПРИЛАДОБУДУВАННЯ

УДК 681.31

МАТЕМАТИЧНЕ ТА ТЕХНІКО-ЕКОНОМІЧНЕ ОБГРУНТУВАННЯ РОЗРОБКИ ІНФОРМАЦІЙНО-ЕФЕКТИВНИХ ОБ'ЄКТНИХ СИСТЕМ КОМП'ЮТЕРНИХ МЕРЕЖ. ЧАСТИНА 1

¹⁾Шевчук Б.М., ²⁾Гераймчук М.Д., ¹⁾Інститут кібернетики ім. В.М. Глушкова НАН України, м. Київ, Україна, ²⁾Національний технічний університет України "Київський політехнічний інститут", м. Київ, Україна

Здійснений аналіз функціонування абонентських (об'єктних) систем комп'ютерних мереж, які забезпечують виконання комплексу операцій багатофункціональної обробки, кодування та формування псевдохаотичних криптистійких і завадостійких пакетів даних мінімальної тривалості, запропоновані підходи до мінімізації інформаційних потоків в комп'ютерних мережах тривалого моніторингу станів об'єктів

Вступ

Бурхливий розвиток комп'ютерних мереж та їх широке застосування в різноманітних галузях людської діяльності зумовлено швидкими темпами розвитку процесорів і мікрокомп'ютерів, а також пакетних мереж передачі інформації. Перспективні напрямки використання комп'ютерних мереж та їх впровадження в промисловість, сільське господарство, охорону здоров'я та навколишнього середовища і таке ін. пов'язані прискореним розвитком безпроводових широко-смугових мереж персонального, локального, регіонального та глобального рівнів зв'язку [1-3], включаючи мережі передачі даних у відповідності до стандартів IEEE 802.xx (Wi-Fi, Wi-Max, Bluetooth, ZigBee та ін.), сенсорні мережі, побудовані у відповідності до технологій Wireless Hart, Eaton RSR, MeshLogic, мережі надширокосмугового зв'язку (UWB-системи), супутникові та мікросупутникові мережі передачі інформації. Основним недоліком широко-смугових мереж, окрім UWB-систем, є використання каналів зв'язку, збиткових по ширині робочої смуги частот, при цьому база каналних сигналів є невеликою і незмінною. Це призводить до неефективного використання каналів зв'язку, а передача даних в умовах високих промислових завад призводить до низької надійності зв'язку. В сучасних UWB-системах обмеженою є дальність зв'язку.

Постановка задачі

В цій статті, з урахуванням розробки безпроводових комп'ютерних мереж при малих початкових та експлуатаційних капіталовкладень, розглядаються питання підвищення швидкості та надійності передачі інформації за рахунок реалізації засобами абонентських (об'єктних) систем методів і алгоритмів зменшення вихідних потоків даних (потоків пакетів даних). Вирішення цієї проблеми досягається шляхом реалізації в місцях виникнення інформаційних потоків

ефективних (швидкодіючих і точних) методів фільтрації і стиску сигналів та зображень, а також зменшення тривалості двійкових послідовностей пакетів даних, що еквівалентно підвищенню коефіцієнта стиску даних без втрат.

Метою статті є обґрунтування побудови абонентських (об'єктних) систем інформаційно-ефективних комп'ютерних мереж, в яких реальна пропускна здатність каналу (підканалів моноканалу) зв'язку наближається до максимально можливої, при цьому пакети даних кодуються і передаються псевдохаотичними інтервальними сигналами, база яких адаптивно змінюється в залежності від поточно визначеного співвідношення сигнал/шум в каналі зв'язку. Основою побудови таких мереж є створення об'єктних систем з підвищеною інформативністю [4], які шляхом реалізації багатофункціональної обробки та кодування первинних потоків даних (сигналів, зображень, відеоданих, двійкових масивів даних) формують та передають крипостійкі та завадостійкі пакети інформації мінімальної тривалості.

Принципи функціонування абонентських (об'єктних) систем інформаційно-ефективних (високоінформативних) комп'ютерних мереж

Функціонування комп'ютерних мереж ґрунтується на пакетній передачі даних абонентськими системами (АС) мережі, які можуть встановлюватись на віддалених об'єктах дослідження і керування (ОДіК) або об'єктах тривалого моніторингу (ОТМ). Враховуючи специфіку та завдання тривалого моніторингу ОДіК(ОТМ) такі АС отримали назву об'єктні системи (ОС). Для підвищення ефективності функціонування комп'ютерних мереж АС(ОС) повинні характеризуватись багатофункціональністю в обробці і кодуванні даних (даних моніторингу) та забезпечувати мінімізацію інформаційних потоків в місцях їх зародження, що в кінцевому рахунку для i -ої АС(ОС), $i = \overline{1, N_{\max}}$, N_{\max} – максимальна кількість активних абонентів комп'ютерної мережі, приводить до зменшення кількості переданих пакетів інформації та їх тривалості. Мінімізація інформаційних потоків на об'єктах перш за все досягається за рахунок виявлення інформативних та достовірних первинних даних на етапі введення вхідних даних від ОДіК(ОТМ) засобами АС(ОС). На основі отриманих даних про якість введеної інформації забезпечуються умови для оптимізації процесів фільтрації і компактного кодування вхідних даних, експрес-аналізу сигналів, зображень, відеоданих та станів ОДіК(ОТМ). З точки зору теорії інформації кожен АС(ОС) можна розглядати як джерело інформації (вихідних пакетів даних), при цьому інформативність вхідних сигналів [5], зображень та результатів обробки даних моніторингу на об'єктах за рахунок реалізації ефективних за обчислювальною складністю і точністю методів і алгоритмів багатофункціональної обробки і кодування даних [4,6,7] перетворюється в інформативність компактних інформаційних кадрів пакетів даних. Відповідно високоінформативні комп'ютерні мережі (ВІКМ), включаючи радіомережі, характеризуються оптимальним вирішенням комплексу проблем в місцях виникнення інформаційних потоків з урахуванням мінімізації сумарного трафіку достовірних даних в спільному ресурсі

ВІКМ – мережі передачі даних [8]. Ефективність функціонування ВІКМ суттєво залежить від оптимізації вирішення проблем введення, обробки, кодування, формування та передачі пакетів даних кожним абонентом мережі. В свою чергу кожна АС(ОС) повинна характеризуватись високою інформативністю при виконанні складових операцій багатофункціональної обробки масивів даних. В результаті програмно-апаратної реалізації обробки та кодування даних в канал зв'язку (моноканал, підканали моноканалу) відправляються оптимізовані по тривалості, криптостійкі та завадостійкі пакети даних. Підвищення інформативності кожної АС(ОС) і ВІКМ в цілому пов'язане з підвищенням достовірності, компактності і захищеності пакетів даних, включаючи захищеність від несанкціонованого доступу до даних і їх підміни, а також захищеність бітів пакетів даних від спотворень каналними завадами. Узагальнений операторний алгоритм функціонування АС(ОС) ВІКМ має вигляд [8]:

$$\begin{aligned}
 \text{ОДіК(ОТМ)} \Rightarrow O_{вд} \Rightarrow O_{фд} \Rightarrow O_{сд} \Rightarrow O_{зд} \Rightarrow O_{зкд} \Rightarrow O_{фІІ} \Rightarrow O_{дкз} \Rightarrow O_{ПерІІ} \\
 \leftarrow O_{фск} \leftarrow O_{ді} \leftarrow O_{ПрІІ}
 \end{aligned}$$

де при передачі даних - $O_{вд}, O_{фд}, O_{зд}, O_{зкд}, O_{фІІ}, O_{дкз}, O_{ПерІІ}$, відповідно оператори введення даних, фільтрації даних, стиску даних, захисту даних, завадостійного кодування даних, формування інформаційних пакетів (ІІ), доступу до каналу зв'язку, передачі ІІ, а при прийомі даних - $O_{ПрІІ}, O_{ді}, O_{фск}$, відповідно оператори прийому ІІ, декодування інформації, формування сигналів керування.

Таким чином АС(ОС) ВІКМ мінімізують вхідні потоки даних та перетворюють їх у псевдохаотичні бітові послідовності мінімальної тривалості, які закодовані з точки зору досягнення попередньо заданої величини ступеня захисту інформації P_z та з урахуванням надійної передачі (доставки) даних по каналах зв'язку з шумами. Тому при введенні інформації (сигналів, фіксованих зображень, відеоданих) АС(ОС) виявляють та компактно кодують найбільш інформативні і достовірні дані, при цьому ефективність багатофункціональної обробки і кодування інформації на АС(ОС) суттєво залежить від вхідного співвідношення сигнал/шум при введенні даних, сумарної величини коефіцієнта стиску даних $K_{см}$, величини P_z , поточного співвідношення сигнал/шум в каналі зв'язку.

Показником ефективності обробки і кодування даних в місцях виникнення інформаційних потоків є відносна величина, яка залежить від тривалості поточних ІІ та визначається виразом

$$\frac{T_{ІІ}^{ноч}}{T_{ІІ}^{\min}} = \frac{t_{сл}^{ноч} + N_b \cdot T_b}{t_{сл}^{\min} + \min\{T_{ІК}\}} \approx \frac{N_b \cdot T_b}{\sum_{i=1}^M f_i \cdot p_i \cdot T_i} \quad (1)$$

де $T_{ІІ}^{ноч}$ – початкова тривалість інформаційного пакету до компактного кодуван-

ня первинних масивів даних; T_{III}^{\min} – мінімізований по тривалості ІІ після компактного кодування даних; $t_{сл}^{поч}$ – тривалість службових даних початкового ІІ (тривалість початку ІІ, поля адреси, поля керування, поля перевірок кодів та кінцевика ІІ); N_b – кількість біт початкового ІІ; T_b – мінімальна тривалість бітового елемента; $t_{сл}^{\min}$ – тривалість службових даних в пакеті, довжиною T_{III}^{\min} ; $\min\{T_{IK}\}$ – мінімальна сумарна тривалість інформаційного кадру після кодування; f_i – частота зустрічі i -ї двійкової послідовності; p_i – номер двійкової послідовності, яка кодується відповідним інтервальним імпульсним сигналом [4]; T_i – тривалість двійкової послідовності; M – максимальна кількість двійкових послідовностей.

Очевидно, що при значних шумах в каналі зв'язку тривалість закодованого інформаційного кадру є значно більшою в порівнянні з аналогічним кадром при використанні “чистого” каналу зв'язку та суттєво залежить від величини бази сигналу, що передається в каналі зв'язку. В свою чергу база сигналу вибирається адаптивно в процесі організації передачі даних між віддаленими абонентами.

Для обробки і кодування об'єктних сигналів, зображень та відеоданих без залучення потужних процесорів, вартісних кодексів і спецпроцесорів актуальними є методи і алгоритми фільтрації-стиску даних, оптимізованих по швидкодії та обчислювальній складності [6,7], які базуються на “сигнальному підході”. Згідно такого підходу в масиві первинних даних вимірювальних сигналів і статичних та динамічних зображень визначаються і кодуються суттєві відліки, точність амплітудно-частотних характеристик яких визначається в процесі реалізації фільтрації-стиску огинаючої кривої сигналів (відеосигналів) та залежить від поточного вхідного співвідношення сигнал/шум $[c/u]_{вх}$ і заданого мінімального значення коефіцієнта стиску даних $K_{сmin}$. Необхідно відмітити, що сам процес фільтрації сигналів і зображень спотворює первинні дані, достовірність відліків яких суттєво залежить від величини $[c/u]_{вх}$. Тому на “зашумлених” ділянках сигналів і зображень молодші двійкові біти відліків є недостовірними і немає змісту точно кодувати вхідні дані. Також на відповідних ділянках огинаючої сигналів (відеосигналів) незначні по амплітуді високочастотні зміни можливо замінити усередненим сигналом. Описані дії дозволяють суттєво підвищити коефіцієнт стиску сигналів і зображень. Відповідно на першому етапі обробки даних вхідні відліки сигналів і зображень необхідно згладити для надійного виявлення суттєвих відліків, якими є екстремуми і точки перегину (точки зміни опуклості кривої). На ділянках огинаючої з високочастотними складовими точками перегину можна нехтувати.

Мінімально необхідна частота дискретизації сигналів вибирається адаптивно в межах величин [9] $f_{\partial K} \leq f_{\partial onm} \leq 2K_{\phi} \cdot f_{max}$, де $f_{\partial K}$ – частота дискретизації сигналу за Котельниковим; $f_{\partial onm}$ – оптимальна частота дискретизації, яка адаптивно підбирається в процесі фільтрації-стиску сигналів; $K_{\phi} \geq 8-10$ – коефіцієнт підвищення частоти дискретизації сигналів $f_{\partial K}$, значення якого суттєво зале-

жить від метрологічних вимог до пристроїв введення і обробки сигналів; f_{\max} – максимальна частота сигналу. Найбільш простим способом адаптивного введення сигналів є їх дискретизація з максимальною частотою $f_{\text{дmax}} = 2K_{\phi} \cdot f_{\max}$ з наступним прорідженням відліків сигналів на величину K_j , де j – поточна величина коефіцієнта прорідження частоти дискретизації. Для простоти кодування значення j вибирається із величин: $j = 1, 2, 3, 4, \dots$; $j = 1, 2, 4, 8, \dots$. З метою зменшення первинних інформаційних потоків і отримання достовірних відліків вимірювальних сигналів доцільно використовувати сігма-дельта АЦП. Оптимальним вирішенням проблем введення сигналів є тактування роботи аналого-цифрового перетворювача з адаптивним періодом кодування $t_k = f(\Delta X_i^{\phi}, [c/w]_{\text{ex}}, K_{\text{cm min}})$, де $\Delta X_i^{\phi} = X_i^{\phi} - X_{i-1}^{\phi}$ – поточне значення приросту відфільтрованого сигналу, обчислене в процесі реалізації оперативної фільтрації, наприклад ковзким способом з мінімальним вікном усереднення $l_y = 4 - 5$ відліків.

З метою мінімізації інформаційних потоків без втрат по точності кодування відліків сигналів на ділянках з високочастотними шумами і на повільно змінних ділянках коефіцієнт прорідження частоти дискретизації сигналів K_j вибирається максимально можливим, а кількість достовірних біт суттєвих відліків при компактному кодуванні $q_0 = 9, 10, 11, 12$ і вибирається в залежності від ступеня "зашумленості" відповідних ділянок сигналів та характеризується різницею $\Delta X_i^u = |X_i^{\phi} - X_i^u|$, де X_i^u – поточний вхідний відлік сигналу з шумами. Шляхом порівняння ΔX_i^u з відповідними пороговими величинами P_1, P_2, \dots, P_s , де s – кількість станів ступеня "зашумленості", опосередковано визначається величина $[c/w]_{\text{ex}}$.

Компактне кодування відліків сигналів ґрунтується на обчисленні величин ΔX_i , ΔX_i^u і виборі величин $f_{\text{донт}}$ і q_0 . На відміну від вимірювальних сигналів первинний потік даних від відеосенсорів визначається роздільною здатністю відеосенсора $N \cdot M$ (N – кількість пікселів в рядку поточного кадру, M – кількість рядків кадру), кількістю біт, вибраних для кодування яскравості пікселя, а також залежить від частоти отримання кадрів і виду зображення (кольорове, монохромне).

Після реалізації алгоритму фільтрації-стиску сигналів (відеосигналів), який передбачає стиск даних з незначними витратами, здійснюється стиск масивів даних без втрат. Компактний масив даних підлягає захисту інформації шляхом гаміювання бітів даних з довготривалими псевдовипадковими послідовностями, які від пакету до пакету є змінними [6]. Кодові ключі генерації довготривалих послідовностей, які підлягають періодичній зміні, фактично є абонентськими секретними кодами. Для захисту даних від підміни та підвищення завадостійкості компактних масивів даних біти завадостійкого кодування випадковим чином маскуються в масиві даних одного або декількох пакетів даних. Основний захист бітових послідовностей пакетів даних від завад здійснюється в про-

цесі формування інтервально-імпульсних сигналів (ПС) пакетів даних [6], при цьому абонент-відправник даних і абонент-адресат (приймач даних) за допомогою центральної станції мікростільника (стільника) або без неї вибирають відповідну базу сигналів, що передаються в каналі зв'язку, та здійснюють надійну передачу даних при дії імпульсних та промислових завад. Слід зазначити, що за рахунок кодування двійкових послідовностей в процесі формування ПС додатково підвищується інформативність АС(ОС) та комп'ютерної мережі в цілому шляхом зменшення тривалості інформаційних кадрів псевдохаотичних інтервальних або шумоподібних пакетів інформацій [4]. Коефіцієнт зменшення тривалості ПП k_g , який відповідає додатковому коефіцієнту стиску двійкових даних в каналі зв'язку, досягає величин $k_g \geq 1.6 - 2$ і більше. Кодування з формуванням ПС досить ефективно може бути використане при побудові перспективних радіомереж з використанням надширококутних сигналів. При цьому бітові інтервали заповнюються надширококутними імпульсними сигналами двох типів: для передачі відповідних "позитивних" (одичних) та "негативних" (нульових) інтервалів. Подальше підвищення інформативності та надійності зв'язку безпроводових мереж досягається внаслідок реалізації абонентами мережі методів надвисокого стиску-захисту даних без втрат та використання технології надширококутної передачі даних в шумах каналу зв'язку.

Висновки

Реалізація інформаційно-ефективної обробки і передачі даних в комп'ютерних мережах досягається за рахунок встановлення в місцях виникнення інформаційних потоків абонентських (об'єктних) систем, які забезпечують виконання комплексу операцій багатofункціональної обробки, кодування та формування псевдохаотичних криптистійких та завадостійких пакетів даних мінімальної тривалості.

Основою багатofункціональної обробки та кодування даних на об'єктах є оптимальний вибір частоти введення даних та кількості достовірних біт суттєвих відліків огинаючої сигналів (відеосигналів). Оперативне компактне кодування сигналів ґрунтується на основі комбінації адаптивного кодування введених даних з урахуванням інформації про вхідне співвідношення сигнал/шум та динамічних показників поточних ділянок сигналів, а також кодування двійкових даних без втрат. Компактні масиви даних підлягають захисту інформації шляхом гаміювання бітів даних з довготривалими псевдохаотичними послідовностями, які від пакету до пакету є змінними. Зменшення тривалості пакетів даних та їх захист від завад в каналі зв'язку здійснюється в процесі формування інтервально-імпульсних сигналів, які можуть передаватись безпосередньо або у вигляді шумоподібних сигналів з адаптивною базою.

Література

1. Шахнович И. В. Современные технологии беспроводной связи, 2-е изд.: - М.: - Техносфера, 2006. – 288 с.

2. Скляр Б. Цифровая связь. Теоретические основы и практическое применение, 2-е изд.: Пер. с англ. – М.: Издательский дом “Вильямс”, 2003. – 1104 с.
3. Ильченко М.Е., Бунин С.Г., Войтер А.П. Сотовые радиосети с коммутацией пакетов. - К.: Наукова думка, 2003. – 266 с.
4. Шевчук Б.М. Теоретичні основи побудови високоінформативних інтелектуальних радіомереж обробки і передачі інформації // Праці міжнар. конф. "Питання оптимізації обчислень (ПОО-ХХІІІ)". – Київ: Ін-т кібернетики ім. В.М. Глушкова НАН України. – 2007. – С. 310-311.
5. Боюн В.П. Динамическая теория информации. Основы и приложения. – К.: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 2001. – 326 с.
6. Шевчук Б.М., Задірака В.К., Фраєр С.В. Ефективні методи фільтрації-стиску та захисту інформації в комп'ютерних мережах тривалого моніторингу станів об'єктів// Штучний інтелект. – 2006. - №3. – С. 804-815.
7. Шевчук Б.М., Зінченко В.П. Оперативна багатофункціональна обробка та передача інформації в моніторингових мережах з використанням мікросупутників// Наукові вісті Національного технічного університету України “Київський політехнічний інститут”. – 2006. - №6. – С. 30-36.
8. Шевчук Б.М. Складові технології побудови високоінформативних комп'ютерних мереж // Матеріали проблемно-наукової міжгалузевої конференції "Інформаційні проблеми комп'ютерних систем, юриспруденції, економіки та моделювання" (ПНМК-2008). – Бучач: Бучачський інститут менеджменту і аудиту, 2008. – С. 44-47.
9. Шевчук Б.М. Методи визначення та відображення показників інформаційних станів об'єктів тривалого моніторингу / Комп'ютерні засоби, мережі та системи. – 2005. - № 4. – С. 78-85.

<p>Шевчук Б.М., Гераїмчук М.Д. Математичне та техніко-економічне обґрунтування розробки інформаційно-ефективних об'єктних систем комп'ютерних мереж. Частина 1.</p> <p>Частина 1 статті посвячена теоретичним аспектам математичного обґрунтування розробки інформаційно-ефективних об'єктних систем комп'ютерних мереж. Показано, що основою багатофункціональної обробки та кодування даних на об'єктах є оптимальний вибір частоти введення даних та кількості достовірних біт суттєвих відліків огинаючої сигналів (відеосигналів). Кодування компактних пакетів даних досягається за рахунок адаптивного стиску даних, гаміювання компактних даних з довготривалими псевдохаотичними послідовностями та формування інтервально-імпульсних сигналів мінімальної тривалості.</p>	<p>Shevchuk B.M., Geraimchuk M.D., Mathematical and techno-economical basis of informational-efficient development Of object system design of computer networks. Part 1.</p> <p>Part 1 . Articles devoted the theoretical aspects of mathematical ground of development of the informational-efficient object systems of computer networks.It is shown that by basis of multifunction treatment and encoding of data on objects there is an optimum choice of frequency of entry of data and amount of reliable bats of the substantial counting out of rounding signals (video-signals). Encoding of compact packages of information is arrived at due to the adaptive compression of data, scaling of compact information with the protracted pseudo - chaotic sequences and forming of interval-impulsive signals of minimum duration.</p>
--	--

*Надійшла до редакції
12 лютого 2009 року*

УДК 681.3.06:519.6

ПРОГНОЗУВАННЯ ТЕХНОЛОГІЧНИХ ПАРАМЕТРІВ МЕТОДАМИ ШТУЧНИХ НЕЙРОННИХ МЕРЕЖ

Пивовар О.М., Вислоух С.П., Національний технічний університет України «Київський політехнічний інститут», м. Київ, Україна

В роботі розглядається питання використання штучних нейронних мереж для прогнозування вихідних технологічних параметрів процесу різання. Наведено приклади прогнозування шорсткості обробленої поверхні та періоду стійкості різального інструменту за допомогою штучних нейронних мереж модуля Neural Network Toolbox системи MatLab

Вступ

Підвищення ефективності приладобудівного виробництва потребує визначення раціональних або оптимальних умов реалізації технологічних процесів. Тому важливою є задача прогнозування вихідних технологічних параметрів за поточними значеннями даних про умови та режими досліджуваного процесу. На сьогодні це здійснюється шляхом математичного моделювання цих параметрів. Питанням математичного моделювання технологічних параметрів процесів обробки деталей на металорізальних верстатах приділяється багато уваги в роботах Г.К. Горанського, М.В. Капустіна, В.Д. Цветкова, В.О. Остаф'єва та багатьох інших вчених в нашій країні та за кордоном. Для отримання математичних моделей вихідних технологічних параметрів процесів різання та їх прогнозування широко використовують планування експериментів і обробка їх результатів методами регресійного аналізу, а також методи кореляційного, дисперсійного аналізу тощо. Але вказані методи потребують великої кількості експериментальних досліджень. Крім того, через велику вартість конструкційних і інструментальних матеріалів, а також енергоносіїв треба використовувати методи моделювання і, відповідно, прогнозування технологічних параметрів, що оснований на незначній кількості дослідів. До таких методів можна віднести евристичні методи, що дозволяють при незначній кількості експериментів отримати якісне прогнозування вихідних технологічних параметрів. Особливе місце серед таких методів має використання штучних нейронних мереж (ШНМ).

Метою даної роботи є вивчення можливостей використання ШНМ для прогнозування вихідних параметрів багатопараметричних систем, до числа яких відносять технологічні процеси обробки деталей різанням.

Розв'язання задачі

Ідея нейронних мереж виникла в результаті спроб змоделювати поведінку живих істот, що сприймають дію зовнішнього середовища та навчаються на власному досвіді. Такого роду ідеї на стику різноманітних областей знання є характерними для науки сьогодення [1, 2]. Цей інтерес до використання ШНМ викликано як теоретичними, так і прикладними досягненнями в даному напрямку. При цьому відкрились можливості використання обчислень в сферах, що

відносились раніш лише до області людського інтелекту, можливості створення машин, здатність яких вчитися та запам'ятовувати нагадує процеси мислення людини, та наповнення новим змістом терміну «штучний інтелект» [2-4].

Аналіз обчислювальних систем, що реалізують ШНМ, показав, що значну гнучкість та функціональність має система MatLab. Завдяки достатньо широкому набору архітектур штучних нейронних мереж та видів їх навчання він дозволяє як будувати ШНМ з довільною архітектурою, так і використовувати широкий спектр цих архітектур, що представлені функціями програмного модуля Neural Network Toolbox цієї системи, та встановлювати метод їх навчання, які дають можливість вирішувати широке коло задач. Тут для створення оптимальної ШНМ можна підібрати необхідні характеристики мережі: вид штучної нейронної мережі, алгоритм та функцію навчання, кількість шарів та число нейронів в кожному окремому шарі й функції активації для кожного шару.

Для навчання мереж програмний модуль дозволяє використовувати функції навчання (TRAINB, TRAINC, TRAINR); градієнтні методи навчання (TRAINGD, TRAINGDA, TRAINGDM, TRAINGDX, TRAINRP); алгоритми методу спряжених градієнтів (TRAINCGF, TRAINCGP, TRAINCGB, TRAINSCG); квазіньютонів алгоритми навчання (TRAINBFG, TRAINOSS, TRAINLM, TRAINBR) тощо.

З метою визначення ефективності використання ШНМ для прогнозування технологічних параметрів проведені експериментальні дослідження процесу токарної обробки сталі 40X13 прохідним різцем із титано-кобальтового твердого сплаву T15K6. При проведенні експерименту оброблювались заготовки зі швидкістю V від 52 до 112 м/хв., глибиною різання t , що варіювалась в межах від 0,25 до 0,5 мм та подачею s в межах 45–90 мм/хв. При цьому вимірювали час різання τ та шорсткість обробленої поверхні R_z .

Результати експериментальних досліджень наведено в таблиці 1.

Таблиця 1. Результати експериментальних досліджень процесу точіння

Вхідні параметри				Вихідна змінна
V , м/хв.	s , мм/хв.	t , мм	τ , хв.	R_z , мкм
112	90	0,5	1,63	40
110	45	0,35	4,93	20
86,4	45	0,25	8,23	10
85,5	70	0,5	10,46	30
67,8	70	0,35	12,69	35
52,4	90	0,5	14,22	80

Оскільки залежність між вхідними та вихідними характеристиками носить явно виражений нелінійний характер, то для навчання була використана штучна нейронна мережа з нелінійними сигмоїдальними функціями активації. Нейронна мережа з прямою передачею сигналу має 2 шари: в першому шарі вико-

ристано 11 нейронів з функцією активації *tansig*, а в другому – 20 нейронів з функцією активації *logsig*.

Навчена таким чином нейронна мережа використана для прогнозування значення вихідного параметра – величини шорсткості обробленої поверхні. Тут в якості вхідних даних взята швидкість різання V – 110 м/хв., подача s – 55 мм/хв., глибина різання t – 0,4 мм та час різання τ – 8 хв. В результаті прогнозування отримано значення шорсткості обробленої поверхні R_z за вказаними режимами різання, що дорівнює 18,3 мкм. При цьому похибка відхилення прогнозованого значення від реального (експериментального) дорівнювала 19%.

Крім того, здійснено прогнозування за допомогою ШНМ величини стійкості різального інструменту при токарній обробці титанового сплаву ВТ6 твердосплавними пластинками із ВК4. В процесі проведення експериментальних досліджень задавали швидкість різання та подачу при постійній глибині різання t , що дорівнювала 1 мм. В якості критерію зношення пластинок вибрано сумарний знос по задній грані 0,9 – 1 мм. Початкові дані для навчання нейронної мережі, що моделює процес, наведено в таблиці 2.

Таблиця 2. Результати експериментальних досліджень процесу токарної обробки титанового сплаву ВТ6

Режими різання		Вихідна змінна
V , м/хв.	s , мм/об	T , хв.
100	0.4	1.8
90	0.4	4.2
70	0.4	7.5
120	0.3	2.7
80	0.3	8.1
160	0.3	3.0
140	0.2	4.5
100	0.2	10.8
90	0.2	16.2

Для навчання штучної нейронної мережі, що моделює процес токарної обробки, використано градієнтний метод.

Прогнозування періоду стійкості інструменту за допомогою навченої ШНМ здійснено при швидкості різання V – 100 м/хв., подачі s – 0,3 мм/об та V – 120 м/хв., подачі s – 0,3 мм/об. При цьому отримано такі значення періоду стійкості інструменту T , відповідно, 5,4 та 5,5 хв., які відрізняються від експериментальних значень не більше ніж на 12%.

Висновки

1 Використання штучних нейронних мереж для прогнозування технологіч-

них параметрів дозволяє в умовах незначної кількості початкової інформації отримати з великою точністю значення вихідної величини досліджуваного процесу.

2. Нейромережеве прогнозування потребує кропіткої роботи з вибору типу і параметрів нейронної мережі та алгоритму її навчання. Але витрати на виконання вказаних робіт значно менші витрат часу та засобів на проведення експериментальних досліджень і прогнозування іншими методами.

3. Штучні нейронні мережі доцільно використовувати як інструментарій для отримання математичних моделей вихідних технологічних параметрів, що дозволить в подальшому здійснювати оптимізаційні розрахунки технологічних процесів.

Література

1. Нейрокомпьютеры и интеллектуальные роботы / Амосов Н.М., Байдык Т.Н., Гольцев А.Д. и др.; под ред. Амосова Н.М. – Киев: Наукова думка, 1991. – 272 с.
2. Neural Network Toolbox User Guide / Beale M., Demuth H. – Natick: Mathworks, 1997. – 700 р.
3. Аведьян Э.Д. Алгоритмы настройки многослойных нейронных сетей // Автоматика и телемеханика. – 1995. – № 4. – С. 106–118.
4. Круглов В.В., Борисов В.В. Искусственные нейронные сети: теория и практика. – М.: Горячая линия – Телеком, 2001. – 382 с.

Пивовар О.Н., Вислоух С.П. **Прогнозирование технологических параметров методами искусственных нейронных сетей.**

В работе рассматриваются вопросы применения искусственных нейронных сетей для прогнозирования выходных технологических параметров процесса резания. Приведены примеры прогнозирования шероховатости обработанной поверхности и периода стойкости режущего инструмента с помощью штучных нейронных сетей модуля Neural Network Toolbox системы MatLab.

Pivovarov O.N., Vislouch S.P. **A prediction of technological parameters by means of artificial neural networks methods.**

Problems of artificial neural networks' application for the prediction of output technological parameters in the time of a cutting process are considered in this work. Examples of the prediction of a processed surface's roughness and a cutting instrument's period of durability by the means of artificial neural networks of the Neural Network Toolbox module of the MatLab system are given.

Надійшла до редакції
14 квітня 2009 року

УДК 004.75, 004.724.2

ОБЧИСЛЮВАЛЬНИЙ КЛАСТЕР ДЛЯ АВТОМАТИЗАЦІЇ НАУКОВИХ ДОСЛІДЖЕНЬ

Порєв Г.В., Національний технічний університет України «Київський політехнічний інститут», м. Київ, Україна

Проаналізовано та обґрунтовано вибір апаратного-програмного забезпечення для побудови

дови обчислювального кластеру. Наведені рекомендації щодо налаштування програмного забезпечення кластеру

Вступ

Певний історичний парадокс полягає в тому, що високопродуктивні комп'ютерні системи (ВПКС) з'явилися раніше персональних, тобто розвиток обчислювальної техніки в ХХ столітті йшов шляхом не тільки мініатюризації схемотехніки, але і децентралізації і персоналізації обчислювальної потужності. Кінцевим наслідком цього процесу стала поява персональних комп'ютерів.

Завдяки розвитку елементної бази сучасні персональні комп'ютери демонструють значно вищу швидкість обчислень і обробки інформації, ніж ВПКС другої половини ХХ століття. Але і в секторі високопродуктивних обчислень за цей період відбулися значні зміни — ВПКС непомітно проникають в повсякденне життя суспільства. Якщо перші справжні ВПКС призначалися для суто військових розробок, то зараз ВПКС керують національними енергетичними та транспортними мережами, вирішують різноманітні наукові математичні та біологічні задачі тощо.

Вважається, що ВПКС відрізняються передусім значною обчислювальною потужністю, яка досягається екстенсивним дизайном обчислювальних модулів, тобто паралельністю обчислень на масиві однотипних обчислювальних блоків. Потрібно відзначити, що персональні комп'ютери і навіть деякі спеціалізовані контролери мають програмні можливості для реалізації паралельних обчислень, але видима «паралельність» в такому випадку досягається швидким переключенням контексту виконуваної задачі в процесорі, який одночасно виконує лише одну послідовність машинних команд.

Сучасні процесори для персональних комп'ютерів, однак, включають в себе апаратні реалізації для справжніх паралельних обчислень, наприклад, включенням більш ніж одного «ядра» (маркетинговий термін для арифметико-логічного пристрою, АЛП) на кристал. Станом на 2009 рік, максимальна кількість ядер в процесорі становить лише 4, тоді як типові ВПКС включають сотні і тисячі вузлів.

Сучасні ВПКС також можуть, і найчастіше включають, апаратно-програмні засоби для реалізації спільної оперативної та спільної довготривалої пам'яті для всіх вузлів, але, на нашу думку, це не є принциповою характеристикою ВПКС в загальному випадку.

Постановка і аналіз задачі

Позиції UNIX як індустріального лідера ринку ВПКС вперше суттєво похитнулися наприкінці 1990-х років, з розширенням функціональності та портабельності операційних систем сімейства GNU/Linux, що дозволило будувати потужні ВПКС на ядрі Linux. Маючи архітектуру, технологічно схожу з UNIX та BSD, ядро Linux, на відміну від них, є повністю безкоштовною для кінцевого користувача системою з відкритим вихідним кодом.

Протягом 10 років архітектура Linux завдяки своїй повній прозорості та зусиллям програмістів з усього світу завоювала практично весь ринок операційних систем для ВПКС. Незважаючи на вихід на ринок ВПКС корпорації Microsoft з операційною системою Windows 2008 for HPC, на сьогодні на 90% ВПКС, побудованих з неспеціалізованих рішень (суто екстенсивний дизайн), встановлена операційна система з сімейства GNU/Linux (рис.1).

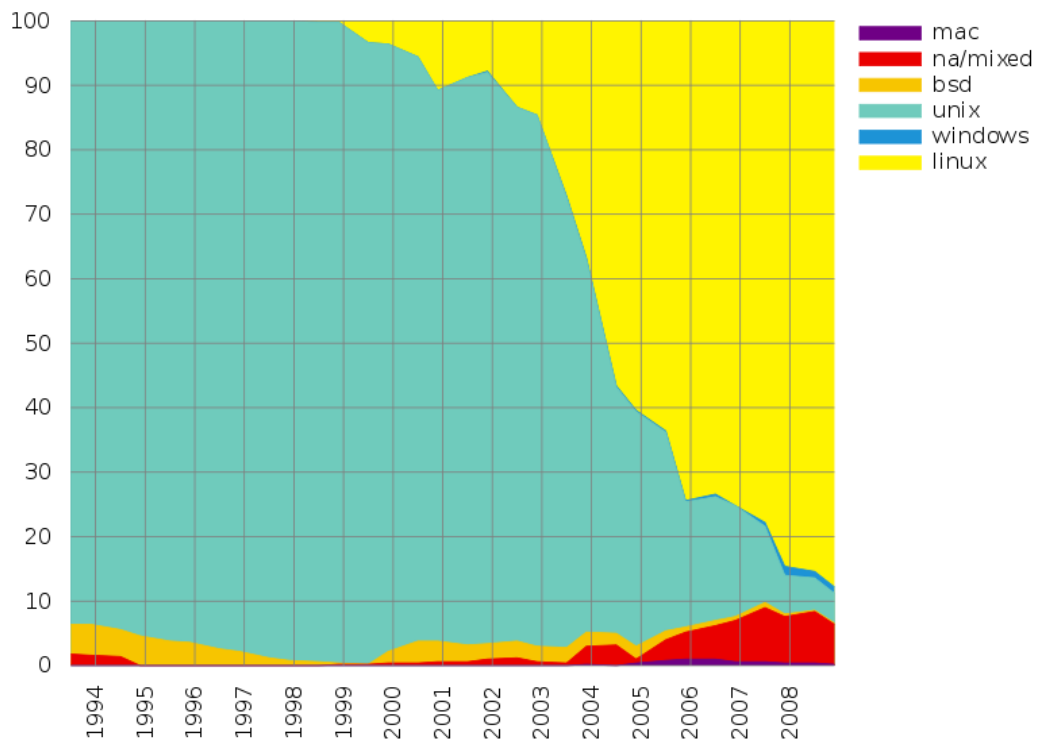


Рис.1. Доля операційних систем на ринку ВПКС в 1994-2008 рр.

Обґрунтування вибору апаратного і програмного забезпечення

За типом комплектуючих, які використовуються для побудови ВПКС, їх можна умовно розділити на «спеціалізований» та «загальний». ВПКС спеціалізованого типу відрізняються тим, що їх апаратна частина в значній мірі проектується і виробляється для вирішення конкретного класу задач. Відповідним чином складається і бюджет проекту, який може сягати сотень тисяч євро. Такі ВПКС реалізують обчислення, специфічні для конкретної задачі апаратно, залишаючи за програмним забезпеченням функції керування. Прикладом ВПКС такого типу можуть бути як одні з перших не програмованих ВПКС так і порівняно недавні розробки комп'ютера Deep Blue — типового прикладу вузькоспеціалізованого ВПКС. Оскільки логіка його ігрової програми, в основному, була побудована на принципі перебору усіх можливих ходів та оцінки їх корисності, то базові одиниці обчислювального алгоритму були реалізовані безпосередньо в його мікропроцесорах. Відтак це потребувало роз-

робки всього програмного забезпечення з урахуванням унікальної специфіки ВПКС. У випадку Deer Blue, як маркетингового ходу корпорації IBM, це було виправдано, оскільки на Deer Blue ніколи не планувалося виконувати інші задачі, і незабаром після завершальних випробувань він був демонтований.

ВПКС загального типу, напроти, збираються з частин та деталей, доступних на користувачькому ринку серверів та їх комплектуючих. Типовий обчислювальний модуль в такому випадку може представляти собою сервер, придбаний у постачальника серверних рішень і встановлений в структуру ВПКС навіть без будь-яких апаратних модифікацій на місці монтажу. Хоча потужні ВПКС реалізують додаткові апаратно-програмні компоненти для прискорення обміну та доступу до даних, вони теж є моделями, що серійно виробляються.

З огляду на застосування ВПКС в наукових дослідженнях важливо те, що не існує формального критерію нижньої межі потужності обчислювальних одиниць. Якщо операційна система може бути встановлена на модулі і підключена в загальну систему до інших модулів, які утворюють ВПКС, то така система формально вважається ВПКС і на ній може бути задіяне відповідне програмне забезпечення.

Для побудови ВПКС (далі — кластер) нами були задіяні 9 обчислювальних модулів, кожен з яких раніше працював як персональний комп'ютер на робочому місці. Апаратні характеристики модулів наведено в таблиці 1.

Таблиця 1. Характеристики модулів обчислювального кластеру

Ім'я вуз-ла	Кількість, тип та частота процесора	Обсяг пам'яті	Ємність диску
Leader	1×AMD K6-2	384 Мбайт	10 ГБайт
Node01	1×Pentium III	256 Мбайт	4 ГБайт
Node02	1×Pentium III	128 Мбайт	4 ГБайт
Node03	1×AMD K6-2	256 Мбайт	8 ГБайт
Node04	1×Pentium II	64 Мбайт	2 ГБайт
Node05	1×AMD K6-2	256 Мбайт	4 ГБайт
Node06	1×AMD K6-2	256 Мбайт	4 ГБайт
Node07	2×Pentium III	512 Мбайт	8 ГБайт
Node08	1×AMD K6-2	64 Мбайт	2 ГБайт

На апаратних ресурсах такого класу, очевидно, неможливо застосовувати ніяких спеціалізованих рішень ринку ВПКС, таких як мережа Infiniband, не

тільки з економічних причин, але і з огляду на те, що на такому кластері не можна розвинути обчислювально-пропускну потужність, яка б потребувала відповідних швидкостей передачі.

Зв'язність обчислювальних модулів кластеру забезпечується мережею Fast Ethernet з пропускнуою спроможністю 100 МБіт/сек і встановленим локально комутатором D-Link.

Вибір базової операційної для кластеру обмежений наступними міркуваннями. Сучасні ВПКС будуть найчастіше на базі ОС GNU/Linux, або, у деяких випадках, на базі ОС FreeBSD або ОС Windows.

Операційні системи сімейства Windows історично не мали штатних можливостей побудови обчислювальних кластерів та ВПКС взагалі, тому всі такі спроби до нещодавна потребували встановлення додаткових сторонніх компонент ПО. Лише з появою ОС Windows Server 2008 HPC Edition операційні системи від Microsoft вийшли на ринок програмного забезпечення для ВПКС. На жаль, навіть мінімальні рекомендовані характеристики системи значно вищі, ніж наявний парк комп'ютерів для даної задачі.

При виборі між GNU/Linux та FreeBSD скористаємося міркуваннями необхідності оптимізації виконуваного коду під конкретну архітектуру процесора а також тією обставиною, що наш кластер є гетерогенним за ознакою такої архітектури — в ньому поєднані процесори різних поколінь та від різних виробників (AMD та Intel). Серед усіх дистрибутивів Linux та BSD штатну, і, більше того, рекомендовану можливість оптимізації під модель процесора має метадистрибутив Gentoo Linux.

Переважна більшість інших дистрибутивів встановлює в систему вже готові, попередньо скомпільовані, орієнтовані на широкий клас процесорів бінарні пакети, тоді як Gentoo Linux надає можливість встановити систему, повністю перекомпільовану на цільовому апаратному забезпеченні з урахуванням не тільки особливостей сімейства процесорів певного виробника, але і конкретної моделі у межах сімейства. Це досягається вказуванням в опціях до компіляторів GCC умовних кодових імен процесорів та вказуванням спеціальних ключів, які можуть задіяти оптимізацію з урахуванням реалізованих в процесорі технологій (таких як MMX, SSE, SSE2 тощо).

Також при виборі дистрибутиву було враховано поточну популярність сімейств операційних систем на користувачькому ринку ВПКС як підґрунтя до наявності широкого вибору програм, широкого загалу користувачів та наявності значного обсягу матеріалу щодо наявних помилок експлуатації та методів їх виправлення.

При проектуванні та початковому налаштуванні програмного забезпечення було враховано досвід фахівців компанії CyberLogic (Квебек, Канада). Проект Adelie Linux R&D цієї компанії є обчислювальним кластером на основі метадистрибутиву Gentoo Linux. Детальний опис процесу побудови так конфігурації кластеру наведено на офіційному веб-сайті метадистрибутиву [1]. Конфігурація кластеру Adelie відрізняється підвищеним акцентом на питаннях безпеки, зок-

рема застосовано офіційні вихідні коди ядра ОС Linux (замість традиційних для Gentoo Linux спеціалізованих ядер), модулі xinetd та iptables для розділення та блокування доступу тощо. В той час як для потужних апаратних засобів та для розташування кластеру в публічному адресному просторі (у випадку проекту Adelie) ці заходи є цілком виправдані, для обчислювального кластеру, який розглядається в даній роботі, потрібно врахувати, що кожен додатковий модуль безпеки займає як дисковий ресурс, так і, що найважливіше, відбиток пам'яті (memory footprint) ядра системи. Що за наведених вище (табл.1) апаратних характеристик може виявитись критичним до швидкодії системи в цілому.

Також необхідно відзначити і ту обставину, що захист систем кластера від зловмисного доступу забезпечується топологічним відділенням його базової мережі від Інтернет, при збереженні можливості доступу ззовні завдяки системі VPN-доступу, встановленому на центральному сервері комп'ютерної мережі.

Можливість оптимізації виконуваного коду всіх компонентів ОС під конкретні сімейство та модель процесора в ОС Gentoo Linux має свої недоліки, основний з яких полягає в тому, що процес перекомпіляції всієї множини пакетів або процес оновлення пакетів до наступних опублікованих версій займає тривалий час. Так, наприклад, монолітне ядро Linux версії 2.6.29 компілюється на одному з вузлів кластера приблизно 20-30 хвилин.

Для прискорення цього процесу нами було встановлено на всіх модулях кластера програмне забезпечення для розподіленої компіляції DistCC. Це, в свою чергу, потребувало встановлення так званих кросс-компіляторів, оскільки в термінології Gentoo Linux деякі модулі кластеру мають архітектуру процесора i586, тоді як інші — i686.

Використання розподіленої компіляції, коли в будь-який момент часу над оновленням або встановленням пакету програмного забезпечення на один модуль працюють всі 9 вузлів дозволило скоротити час оновлення в середньому в 3-4 рази. Невідповідність зумовлена тим, що розподіляється по кластеру тільки безпосередній процес компіляції, а попередня конфігурація та обробка проводиться на цільовому модулі.

Для централізованого керування модулями кластеру було створено сценарій, вихідний код якого наведений нижче

```
#!/bin/bash
#
# sixth node temporarily excluded
#
nodes="leader node01 node02 node03 node04 node05 node06 node07 node08"
#
for node in ${nodes}; do {
echo === RUNNING COMMAND ON "$node" ===
ssh -l root $node "$1"
}; done;
#
```

Даний сценарій, написаний в синтаксисі командного інтерпретатора (оболонки) bash, реалізує послідовне виконання заданої команди на всіх модулях

кластеру. Отже, без доступу до керуючої консолі кожного модуля досягається одноманітність встановлених версій програмного забезпечення та конфігурації модулів, що критично для однорідної роботи обчислювального кластеру.

Для виконання безпосередніх розподілених обчислень на всі модулі кластеру було встановлено бібліотеку OpenMPI, яка реалізує стандарт обміну даними в паралельних обчислювальних системах. Оскільки не планується надавати одночасний (конкурентний) доступ до обчислювальних ресурсів кластеру, системі планування задач, таку як Torque Resource Manager встановлено не було.

Для успішної роботи OpenMPI з використанням мережної файлової системи NFS налаштовано спільний ресурс, який вказує на домашній каталог виділеного користувача, та внесено ключі SSH-авторизації для уникнення необхідності багаторазового вводу паролю під час виконання розподілених задач.

Висновки

Проаналізовано та обґрунтовано вибір апаратного та програмного забезпечення для побудови обчислювального кластеру для автоматизації наукових досліджень. Такий кластер побудовано з використанням тих же технологій, які використовуються для створення потужних обчислювальних кластерних систем промислового та наукового значення.

Наведені в статті рекомендації щодо конкретних пакетів програмного забезпечення та їх конфігурації надають можливість впровадження аналогічних кластерів на кафедрах високотехнологічних спеціальностей, що, в свою чергу, сприятиме підвищенню рівня професійної підготовки студентів.

У порівнянні з аналогами даний обчислювальний кластер має менші апаратні вимоги, вищу ступінь адаптації до гетерогенного апаратного забезпечення, і, як наслідок, вищу продуктивність.

Література

1. High Performance Computing on Gentoo Linux: <http://www.gentoo.org/doc/en/hpc-howto.xml>

<p>Порев Г.В. Вычислительный кластер для автоматизации научных исследований Проанализирован и обоснован выбор аппаратно-программного обеспечения для построения вычислительного кластера. Приведены рекомендации насчёт настройки программного обеспечения кластера.</p>	<p>Poryev G. V. Computing Cluster to aid the Scientific Research The choice for the software and hardware used to build the computing cluster was analyzed and substantiated. The recommendations concerning the software package and setup was outlined.</p>
---	--

*Надійшла до редакції
12 березня 2009 року*

УДК 519.85 УДК 531.7

МОЖЛИВІСТЬ ВИЗНАЧЕННЯ КОНТУРІВ ДЕТАЛІ ПРИ ОБРОБЦІ ЗОБРАЖЕННЯ З ВИКОРИСТАННЯМ СИСТЕМ РОБОТО-ТЕХНІЧНОГО ЗОРУ

Борковський О.В., Національний авіаційний університет, м. Київ, Україна

У статті розглянуто задачу обробки візуальної інформації при роботі з просторовими об'єктами, у тому числі при вимірюванні геометричних параметрів об'єктів. Розроблено алгоритм ідентифікації деталей

Вступ

Проблема аналізу зображення деталей в промисловості при контролі деталей, виробів представляє особливий інтерес. За допомогою сучасних систем технічного зору, сучасної системи реєстрації, розпізнавання та ідентифікації зображень можна контролювати у виробництві точність та відповідність геометричних розмірів деталей та дефекти деталей.

Аналіз останніх досліджень та публікацій

З аналітичного огляду інформаційних джерел, стають зрозумілими останні тенденції розвитку в даній галузі. А саме – найсучасніші прилади подібної та такої самої специфіки мають за основу лазерний сканер, інформація про сканування передається у комп'ютер. Обробка отриманої таким способом інформації проводиться апаратно-програмними методами [1-3]. Одиночне зображення невідомого об'єкту фіксується на деякому носії сприймаючого (чутливого) елементу розпізнаючого комплексу (фотокамера, відеокамера). При вводі у ЕОМ це зображення об'єкту дискретизується, квантується за яскравістю та його елементи дискретизації кодуються.

Одночасно з метою ущільнення отриманої інформації, диференціювання за значимістю для подальшого аналізу її мінімізують електронним або алгоритмічним шляхом при ідентифікації використовується лише зовнішній контур зображення[2-4].

Паралельно проводиться аналіз зображення з метою виявлення на ньому окремих опорних елементів, необхідних для проведення нелінійних перетворень зовнішнього контуру зображення об'єкту.. Далі об'єкт ідентифікується та при цьому або відбувається класифікація невідомого об'єкту, або, якщо ПК не дає відповіді, надається запит нової реалізації об'єкту, знятої під іншим ракурсом.

З метою автоматизації завдань розпізнавання та обробки візуальної інформації, вже багато десятиліть ведуться розробки систем технічного зору, що знаходять своє застосування у робототехнічних комплексах (в тому числі координатно-вимірювальних машинах) різного призначення [3-7].

Постановка задачі

Провести вимірювання деталі за допомогою робототехнічного зору швидкісною відеокамерою.

Розв'язання поставленої задачі

В режимі реального часу проведемо вимірювання деталі з визначенням відхилення від форми та розташування поверхонь це дуже узагальнена схема роботи, що не враховує подробиці, пов'язані з налагодженням системи, принципом синхронізації, алгоритмом аналізу та інше. При необхідності перевірки візуальних характеристик предметів на виробничих лініях така система машинного зору може дати відмінні результати.

Розглянемо завдання кодування зображень, що складаються із плоских фігур. На рецепторному полі фігури представляють безліч приналежних їм крапок, що утворюють зв'язну область. Крапка b_i , що перебуває в цій області, називається граничною, якщо хоча б одна із чотирьох сусідніх крапок рецепторного поля зображенню фігури не належить. Кожна фігура задається множиною $B\{b_1, b_2, \dots, b_r\}$ граничних крапок, які у сукупності утворюють її контур. Від способу подання контуру фігур залежить компактність подання зображення та ефективність обробки. Контур апроксимується ламаною лінією або задається параметрично.

Фігура представляється у вигляді багатокутника, що задається безліччю вершин, які називаються вузлами апроксимації. Потрібно побудувати такий багатокутник, що містив би мінімальне число вершин і в той же час представляв фігуру із заданою точністю. Вимоги до точності апроксимації сформулюємо в такий спосіб: відстань від кожної крапки $b_i \in B$ до найближчої крапки перетинання відрізка апроксимуючої ламаної з деякою лінією растра, що проходить через b_i , не повинна перевищувати кроку рецепторної сітки. Для конкретності покладемо, що вузли апроксимації повинні вибиратися з множини граничних крапок.

Розглянемо застосування швидкісної відеокамери у системах технічного зору. Автономна система машинного зору на базі швидкісної відеокамери FastVideo з DSP-процесором працює в такий спосіб.

1. Відеосигнал деякого періодичного процесу реєструється матрицею відеокамери (наприклад швидкісна відеокамера FastVideo-250, FastVideo-500 або швидкісна камера FV-250DSP із вбудованим DSP-процесором BlackFin чи високошвидкісна камера FV-500DSP із вбудованим DSP-процесором TigerSharc).

2. З матриці дані попадають у контролер відеокамери (наприклад контролер на базі процесора TigerSharc).

3. З контролеру камери дані передаються в контролер DSP-процесору.

При побудові кожного відрізка ламаної доцільно дотримуватися наступних положень. Якщо відомо початкову крапку $b_i \in B$, то як кінець відрізка вибирається максимально віддалена від b_i крапка b_{i+k} , така, що відстань від кожної крапки контуру, розташованої між b_i та b_{i+k} , до відрізка $b_i b_{i+k}$ задовольняє

сформульованим вище вимогам по точності апроксимації. Для визначення такої максимально віддаленої крапки b_{i+k} можна скористатися наступним алгоритмом.

Помістимо початок координат (x_0, y_0) у відому початкову крапку b_i . Зв'яжемо з кожною крапкою $(x_j, y_j) \in B$ деякі обмеження на область, у якій може лежати черговий вузол апроксимації. Для цього проведемо з початку координат через сусідні з (x_j, y_j) крапки растру чотири промені. Виберемо два w_j і W_j , так, щоби утворений ними кут мав максимальне значення. Очевидно, що черговий вузол апроксимації повинен належати області w_j , обмеженої цими променями.

Визначення кінцевої крапки b_{i+k} здійснюється перебором крапок b_{i+1}, b_{i+2}, \dots в напрямку обходу контуру. Нехай на черговому кроці перебору розглядається крапка b_j , і, в результаті попередніх кроків алгоритму, визначена область Ω_{j-1} , якій повинен належати вузол апроксимації. Тоді при розгляді крапки b_j , ця область коректується за наступним правилом: $\Omega_{j-1} = \Omega_{j-1} \cap w_j$. Якщо $b_j \in \Omega_{j-1}$, то мається на увазі, що $b_{i+k} = b_j$. В противному випадку значення b_{i+k} не міняється. На початку реалізації алгоритму покладається, що Ω_0 дорівнює всій площині.

Вважається, що відрізок побудований, якщо на черговому кроці алгоритму область Ω_j порожня. Побудова відрізка ламаної ілюструється рис. 1. Контурні крапки помічені числами 0, 1, 2... (крапка 0 — початкова). На рисунку зображені промені, що обмежують область Ω_j . Номер j променю w_j або W_j дорівнює номеру тієї вершини, що вносить відповідне обмеження на область Ω_j . Промені мають кілька міток, якщо обмеження на область Ω_j , внесені декількома крапками співпадають. Область Ω_6 представляє бісектрису кута xOy , а область Ω_7 порожня.

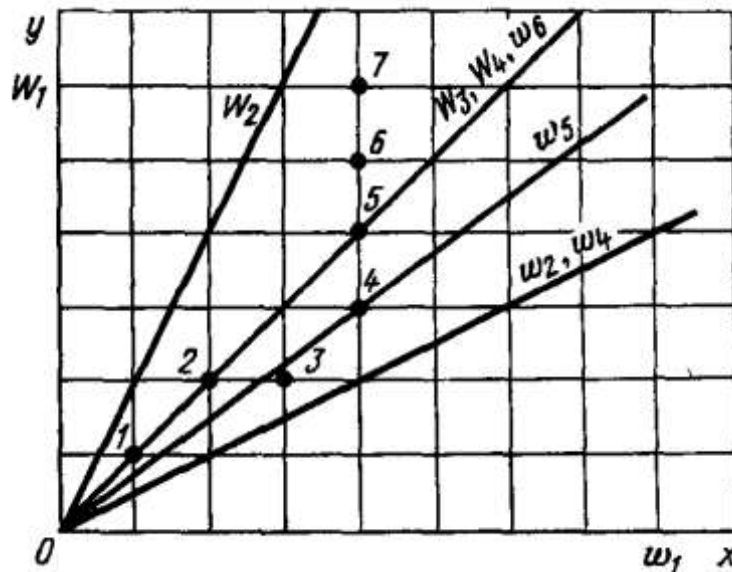


Рис.1. Побудова відрізка ламаної

Апроксимація контурів фігур заданими кривими

Перевагами володіє подання аналізованих фігур у вигляді деяких стандартних, добре вивчених у геометрії фігур, близьких до апроксимованих за формою та розмірами[6]. Таке подання відрізняється компактністю та дозволяє застосувати для обробки зображень апарат аналітичної геометрії. Розглянемо завдання апроксимації контурів фігур у наступній постановці. Нехай заданий дискретний замкнений контур $B = \{b_1, b_2, \dots, b_i, \dots, b_r\}$ з координатами контурних крапок $b_i = (x_i, y_i)$ та множина стандартних фігур, кожна з яких визначається замкненою кривою (ламаною), що описує її границю.

Множину кривих позначимо через Φ . Потрібно вибрати $\phi \in \Phi$, для якої сума квадратів відстаней до крапок контуру мінімальна. Алгоритм апроксимації будемо ілюструвати на прикладі множини Φ , що складається із прямокутника, еліпса, ромба та рівнобедреного трикутника, в якому кут між двома рівними сторонами не перевищує 60° .

Зафіксуємо деяке канонічне положення кривих щодо осей декартової системи координат. З цією метою побудуємо для кожної кривої вісь, що проходить через її центр ваги, щодо якої момент інерції кривої приймає мінімальне значення. Сумістимо вісь x із цією головною віссю ваги, а центр системи координат - із центром ваги. Канонічне положення кривих ілюструється рис. 2.

Для трикутника можливі два канонічних положення, що виходять один з другого поворотом трикутника на кут 180° .

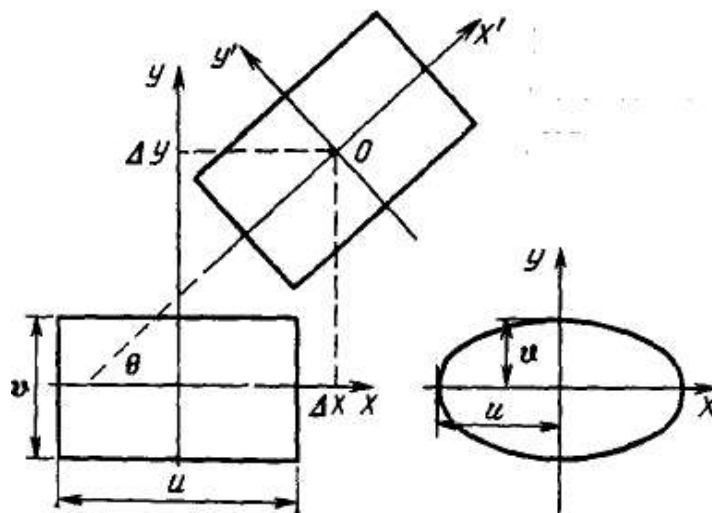


Рис. 2. Канонічне положення прямокутника, еліпса, ромба та трикутника відносно осей декартової системи координат

Зручно розрізняти ці два подання як дві різні фігури. Нехай у множину Φ входить одна з них, а саме та, котра зображена на рисунку. Очевидно, будь-яке розташування кривої в растрі може бути зведене до канонічного переносу центру системи координат по осі x та y на відстані Δx і Δy і повороту осей координат на кут θ . Прийmemo також, що в канонічній системі координат

апроксимуючі криві однозначно визначаються двома параметрами, які позначимо через u та v . Для кривих множини Φ ці параметри мають наступний сенс: у прямокутнику – це довжини сторін, в еліпсі – велика та мала півосі, у ромбі – довжини діагоналей, а трикутнику – основа та висота. Завдання побудови кожної кривої $\phi \in \Phi$, таким чином може бути зведена до відшукування компонентів вектору $G = (\Delta x, \Delta y, \theta, u, v)$.

В цілому алгоритм апроксимації розбивається на три етапи. На першому з них вирішується завдання суміщення контуру із кривими множини Φ . При цьому визначаються Δx , Δy та θ і перетворюється система координат, в якій визначені контурні крапки. Таке перетворення є загальним для суміщення з будь-якою кривою $\phi \in \Phi$. На другому етапі для кожної кривої знаходять її геометричні параметри u та v . Третій етап зводиться до вибору з множини побудованих на другому етапі кривих такої, сума квадратів відстаней якої від контурних крапок мінімальна. Результати реалізації кожного з етапів можуть становити самостійний інтерес. Наприклад, параметри u та v усіх побудованих на другому етапі алгоритму кривих можуть бути використані як ознаки в завданнях класифікації.

Головна вісь ваги дискретного контуру визначається як пряма, сума квадратів відстаней від якої до крапок контуру мінімальна. Координати x_0 та y_0 центру ваги обчислюють за формулами: $x_0 = \sum_{i=1}^r x_i / r$; $y_0 = \sum_{i=1}^r y_i / r$. Рівняння головної осі ваги представляється у вигляді: $(y - y_0) = k(x - x_0)$.

З умови мінімуму суми квадратів відстаней від крапок контуру визначається кутовий коефіцієнт k :

$$k = \frac{\sum_{i=1}^r [(x_i - x_0)^2 + (y_i - y_0)^2]}{2 \sum_{i=1}^r (x_i - x_0)(y_i - y_0)} \pm \sqrt{\left[\frac{\sum_{i=1}^r [(x_i - x_0)^2 + (y_i - y_0)^2]}{2 \sum_{i=1}^r (x_i - x_0)(y_i - y_0)} \right]^2 + 1}$$

Поміщаємо центр нової системи координат у крапку (x_0, y_0) та суміщаємо вісь абсцис із головною віссю ваги. Координати крапок контуру x'_i і y'_i у новій системі координат обчислюють за формулами:

$$x'_i = [k(x_i - x_0) - (y_i - y_0)] / \sqrt{1 + k^2}; \quad y'_i = [(x_i - x_0) - k(y_i - y_0)] / \sqrt{1 + k^2}.$$

Для визначення параметрів u та v застосуємо штучний прийом, що зводиться до розбивки множини контурних крапок і кривих ϕ на частини.

Покладемо, що вихідний контур розпадається, як показано на Рис. 3, на дві пари замкнених контурів, утворених відповідно крапками контуру, що лежать у верхній, нижній, лівій та правій напівплощинах, а також крапками перетину осей координат з лініями координатної сітки. Всі перетворення роблять у новій системі координат. Позначимо через δx середню відстань від осі y у центрів ваги двох контурів, розташованих у лівій та правій напівплощинах. Аналогічно через δy позначимо такі відстані від осі x для контурів, розташованих у верхній і

нижній напівплощинах. Величини δx і δy можуть бути визначені із співвідношень:

$$\delta x = \sum_{i=1}^r \frac{|x_i'|}{r + 2(y_B - y_H) - 2}; \quad \delta y = \sum_{i=1}^r \frac{|x_i'|}{r + 2(x_{II} - x_{II}) - 2}$$

де x_{II} , x_{II} , y_B , y_H — відповідні координати крапок перетинання контуру з осями координат. Параметри u та v можуть бути визначені, виходячи зі значень δx та δy для кожної із кривих, що належать множині ϕ . Так в апроксимуючому прямокутнику величини δx та δy , знайдені тим самим способом, що і для вихідного контуру, визначаються співвідношеннями:

$$\delta x = u_{II}/4; \quad \delta y = v_{II}/4,$$

де u_{II} та v_{II} - довжини сторін прямокутника. Звідси $u_{II} = 4\delta x; v_{II} = 4\delta y$.

Співвідношення, що визначають розміри великої u_E та малої v_E півосей еліпсу визначаються: $u_E = \frac{3}{4}\pi\delta x; v_E = \frac{3}{4}\pi\delta y$.

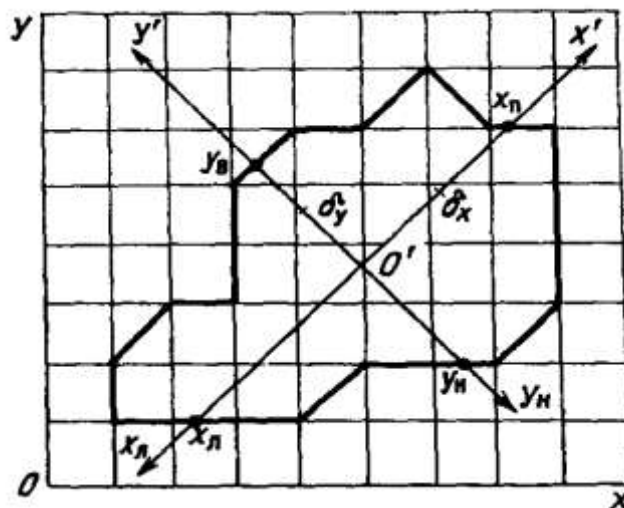


Рис. 3. Визначення параметрів апроксимуючих кривих

Аналогічно обчислюються діагоналі u_P та v_P ромбу, а також основа u_T та висота v_T трикутника:

$$u_P = 6\delta x; \quad v_P = 6\delta y, \quad u_T = 6\delta y, \quad v_T = (17/81)^{-1} \delta x \approx 4.76\delta x.$$

На третьому етапі алгоритму кожній кривій $\phi \in \Phi$ співставляється оцінка D_ϕ , що обчислюється за формулою: $D_\phi = \sum_{i=1}^r d_\phi^2(x_i', y_i')$, де $d_\phi(x_i', y_i')$ - відстань від крапки контуру (x_i', y_i') до кривої ϕ . З множини кривих в якості апроксимуючої вибирають криву з мінімальним значенням оцінки. Позначимо через d_{II} , d_E , d_P та d_T відстань у перетвореній системі координат від деякої крапки (x, y) до прямокутника, еліпса, ромба та трикутника відповідно. Значення d_{II} , d_P та d_T обчислюють за формулами:

$$d_{II} = \min \left\{ \left| x + \frac{u_{II}}{2} \right|, \left| x - \frac{u_{II}}{2} \right|, \left| y + \frac{v_{II}}{2} \right|, \left| y - \frac{v_{II}}{2} \right| \right\}; d_p = \min \left\{ \left| \frac{2v_p x + 2u_p y - u_p v_p}{l_p} \right|, \left| \frac{-2v_p x + 2u_p y - u_p v_p}{l_p} \right|, \left| \frac{2u_p x - 2v_p y - u_p v_p}{l_p} \right|, \left| \frac{-2u_p x - 2v_p y - u_p v_p}{l_p} \right| \right\}; l_p = 2\sqrt{u_p^2 + v_p^2};$$

$$d_T = \min \left\{ \left| x + \frac{v_T}{3} \right|, \left| \frac{u_T x + y - \frac{v_T}{3}}{l_T} \right|, \left| \frac{-\frac{u_T x}{2v_T} + y + \frac{v_T}{3}}{l_T} \right| \right\}; l_T = \sqrt{\frac{u_T^2}{4v_T^2} + 1}.$$

Обчислення точних значень параметру d_E для еліпса пов'язано із значними аналітичними труднощами. Тому d_E визначають приблизно, як половину модуля різниці між сумою відстаней від 1-ї крапки контуру до фокусів еліпса та константою еліпса $2u_E$:

$$d_E = \frac{1}{2} \left| \sqrt{(y_i')^2 + (x_i' - l_E)^2} + \sqrt{(y_i')^2 + (x_i' + l_E)^2} - 2u_E \right|,$$

де l_E – половина фокусної відстані еліпса.

На рис. 4. представлені приклади виконання програми для двох варіантів вихідного контуру. В якості апроксимуючої кривої обрані: a – еліпс; b – прямокутник.

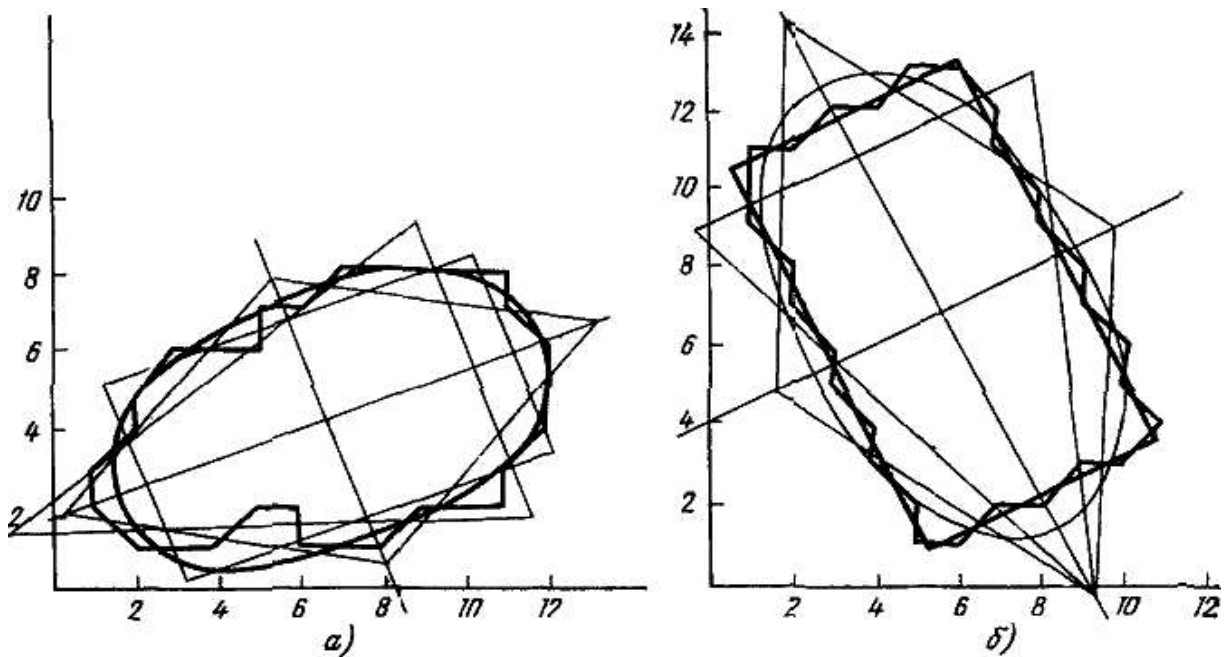


Рис. 4. Приклад параметричної апроксимації контуру фігур

Висновки

Запропонований алгоритм може бути застосований для опису контурів зображень широкого класу об'єктів. На перших двох етапах алгоритму як міра близькості крапок контуру та апроксимуючої кривої використовується не сума

квадратів відповідних відстаней, а співпадіння деяких характерних ліній і крапок. Показано що, при проектуванні основних вузлів систем технічного зору в якості елемента реєстрації може бути використана швидкісна відеокамера.

У подальшому необхідно вирішення задачі оптимізації точності вимірювання за допомогою робото-технічної системи.

Література

- 1.Идентификация визуальных объектов с использованием знаний/ Д.А. Денисов, А.К. Дудкин// Системы технического зрения. – М. Наука, 1991. – 200с.
- 2.Ерганжиев Н.А. Цветное телевидение в измерительной технике. – М.: Связь, 1980. – 104 с.
- 3.Абламейко С.В., Лагуновский Д.М. Обработка изображений: технология, методы, применение. – Минск: Институт технической кибернетики НАН Беларуси, 1999. – 300 с.
- 4.Горелик С.Л., Кац Б.М., Киврин В.И. Телевизионные измерительные системы. – М.: Связь, 1980. – 168 с.
- 5.Гринберг А.Д., Гринберг С. Цифровые изображения: Практическое руководство: Пер. с англ. – Минск: Попурри, 1997. – 400 с.
- 6.Гонсалес Р., Вудс Р. Цифровая обработка изображений М.:изд-во Техносфера, 2006.-1701 с.
- 7.Гуленко И.Е. Система видеозахвата и анализа движения – распознавание трансформаций и движения объекта. – Труды конференции “Новые информационные технологии” (Судак, Крым, 15–25 мая 2004 г.), с. 141-142

<p>Борковский А.В. Возможность определения контуров деталей при обработке изображений с использованием систем робото-технического зрения</p> <p>В статье рассмотрена задача обработки визуальной информации при работе с пространственными объектами, в том числе при измерении геометрических параметров объектов. Разработан алгоритм идентификации деталей.</p>	<p>Borkovskiy A.V.The possibilities of the determination of the detail's contour by it treatment with using of robotics vision systems.</p> <p>In work state the processing of visual information at working with the dimensional objects, including at the measurement of geometrical parameters of objects is offered. The algorithm of identification of details are created.</p>
---	---

Надійшла до редакції
3 березня 2009 року

МЕТОД ВИПАДКОВОГО ПОШУКУ ОПТИМУМУ ФУНКЦІЇ ЗІ ЗМІННИМ РАДІУСОМ ПОШУКУ

Левандовська І.В., Національний технічний університет України «Київський політехнічний інститут», м. Київ, Україна

В статті наведено спосіб знаходження оптимального значення функції за допомогою методу випадкового пошуку зі змінним радіусом пошуку та випадковим напрямком. Показано переваги запропонованого оптимізаційного алгоритму визначення екстремальних значень функцій та приклад його застосування

Вступ

Процес оптимізації лежить в основі усієї інженерної діяльності, оскільки

функції фахівця полягають у тому, щоб, по-перше, проектувати нові більш ефективні та менш дорогі технічні системи, та, по-друге, розробляти методи підвищення якості функціонування існуючих систем. В практичній діяльності часто з багатьох можливих рішень задачі необхідно вибрати оптимальне. Для знаходження «кращого» значення існує ряд методів, що покладені в основі математичного моделювання.

Наразі існує багато методів оптимізації функції. Один з них – метод випадкового пошуку з постійним радіусом пошуку та випадковим напрямком [1]. Суть методу полягає в тому, що пошук оптимального значення проводиться на поверхні сфери (або кола) з певним радіусом у довільному напрямку.

При «вдалому» виборі точки радіус сфери збільшується, що призводить наближення до шуканої точки екстремуму. Але при збільшенні радіуса сфери збільшується площа поверхні, з якої довільно обираються точки для дослідження, що зменшує точність обчислень, збільшує кількість машинного часу, необхідного для розв'язання задачі, і, якщо не застосувати певні математичні закони для вибору точок дослідження, то це може привести до хибного результату. Тому виникла потреба у створенні методу, який не має вказаних недоліків, є простим для реалізації та має більшу точність і швидкодню, за рахунок зменшення площі вибору точок дослідження та врахування характеру їх вибору.

Розробка методу оптимізації

Згідно з запропонованим методом формується коло (сфера або гіперсфера), на контурі якого довільно вибирається певна кількість точок. Шляхом аналізу вибраних точок визначається напрямок наступного кроку пошуку. При знаходженні області, в середині якої розміщено оптимум функції, відбувається уточнення точки екстремуму до отримання відповіді з заздалегідь зазначеною точністю обрахунків ϵ .

Визначення оптимального значення функції згідно з запропонованим методом виконується в такій послідовності:

1. Задається початкова точка x_0 . x_0 – центр кола (якщо пошук проводиться у двовимірному просторі) або сфери (у тривимірному), гіперсфери (у багатовимірному просторі).

2. На поверхні круга або сфери (в залежності від кількості змінних) довільно вибирається певна кількість точок. Визначається значення функції у кожній з цих точок. Точку, у якій значення функції буде найменшим та рівним або менше ніж значення функції в точці x_0 , позначається \tilde{x}_0 .

3. Через точки x_0 і \tilde{x}_0 проводиться пряма/площина та перевіряється значення функції в кожній точці цієї прямої з кроком α . α – невелике число.

4. Точка, що знаходиться на прямій $x_0\tilde{x}_0$ та у якій функція має найменше значення, позначається як x_1 . Відбувається перехід до нового кроку ітерацій, де x_1 – центр кола або сфери.

Надалі пошук здійснюється згідно з пунктами 2, 3, 4. На n -ітерації точка x_n являється центром кола або сфери, а \tilde{x}_n є випадковою точкою, що знаходиться на поверхні фігури, в якій функція має найменше значення. Надалі здійснюється пошук найменшого значення вздовж прямої $x_n\tilde{x}_n$. Знайдену точку, у якій функція має найменше значення, позначають як x_{n+1} та відбувається перехід до $n+1$ -ітерації.

5. Якщо у ході пошуку \tilde{x}_n не було знайдено жодної точки, що б відповідала поставленим вимогам, необхідно повторити пошук знов розпочинаючи з пункту 2. Щоб підтвердити, те що в середині даного кола або сфери знаходиться оптимальне значення функції, необхідно повторити пошук \tilde{x}_n m разів. m – довільно вибране ціле число.

6. При виявленні області, в якій знаходиться оптимальне значення функції на k -ітерації, фіксується точку x_k , як центр, та зменшується радіус кола/сфери на величину β .

7. Обирається певна кількість довільних точок, що знаходяться на поверхні кола або сфери, і знаходиться значення функції у цих точках.

8. Якщо жодне значення функції в випадково вибраних точках не являється меншим ніж значення функції у точці x_k , то радіус кола (сфери) зменшується на величину β , і знов повторюються дії зазначені в пункті 6.

Якщо значення будь-якої довільно вибраної точки менше ніж значення функції в точці x_k , то центр кола або сфери переміщується у дану точку і повторюються дії зазначені у пункті 7.

9. Пошук відбувається до тих пір, доки радіус кола (сфери) не буде меншим або дорівнює точності обрахунків ϵ .

Блок-схема алгоритму знаходження оптимального значення функції за допомогою даного методу зображені на рис. 1.

Визначимо мінімум функції $f(x_1, x_2) = (x_1 - 4)^2 + (x_2 - 3)^2 + 3$ за допомогою розробленого методу. Для цього виберемо початкову точку з координатами $(-3; -2)$ і проведемо сферу з одиничним радіусом.

Виберемо 5 довільних точок, що лежать на поверхні сфери. У точці з координатами $(-2.3; -1.2)$ функція буде мати значення $f(x_1, x_2) = 60.3$, що являється найменшим значенням серед вибраних точок. Через точки з координатами $(-3; -2)$ та $(-2.3; -1.2)$ проведемо площину паралельну осі Oz.

Перебираючи значення функцій в точках, що належать площині, знайдемо точку, в якій функція матиме найменше значення – $(1.2; 4.6)$, $f(x_1, x_2) = 13.4$. Проведемо сферу з одиничним радіусом та центром в точці з координатами $(1.2; 4.6)$. Довільно виберемо 5 точок на поверхні сфери.

Найменше значення функції серед обраних точок знаходитиметься в точці з координатами $(2.3; 4.4)$, $f(x_1, x_2) = 7.85$. Через точки з координатами $(1.2; 4.6)$ та $(2.3; 4.4)$ проведемо площину паралельну осі Oz.

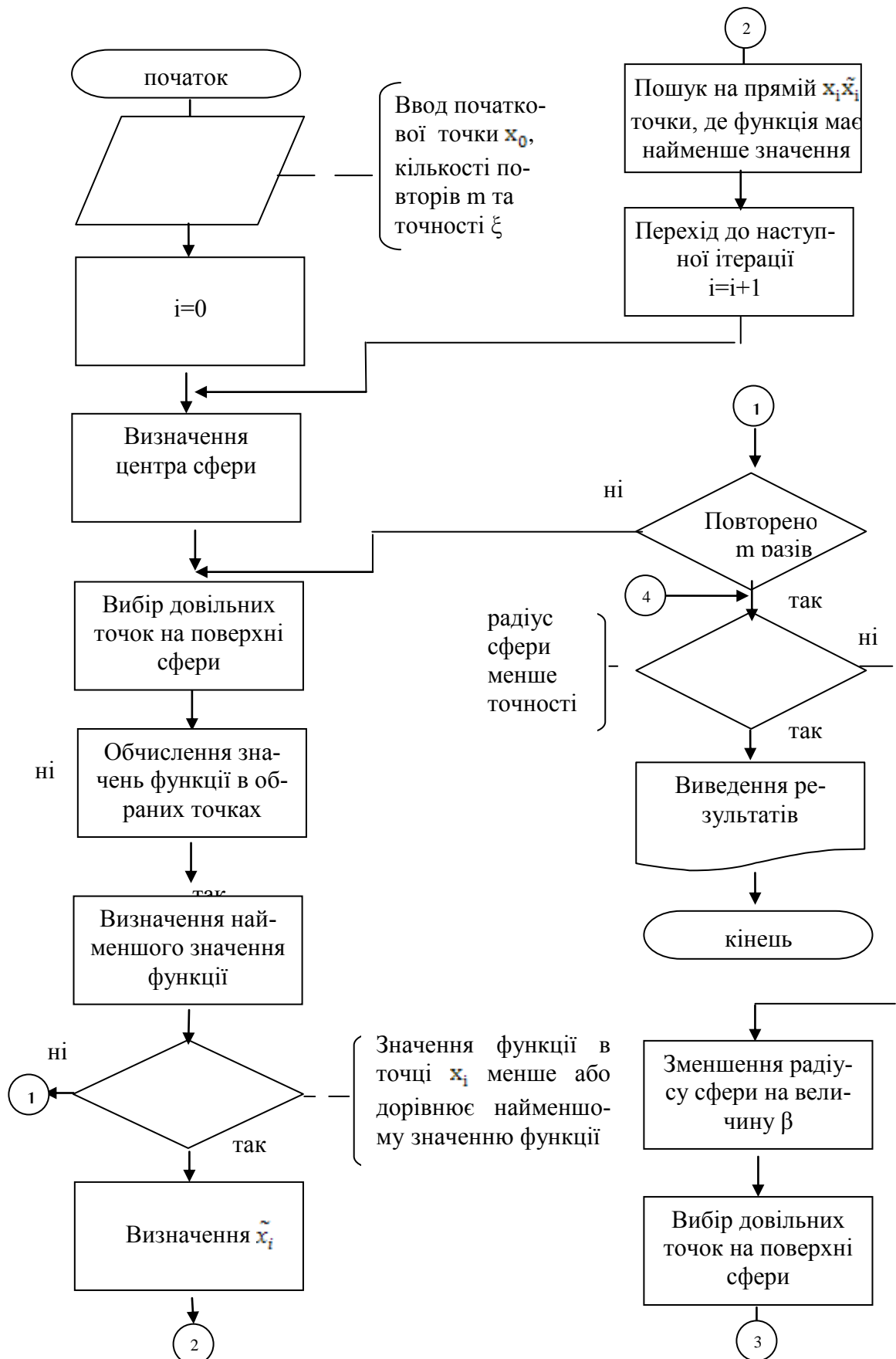


Рис. 1. Алгоритм знаходження мінімуму функції за допомогою розробленого методу.

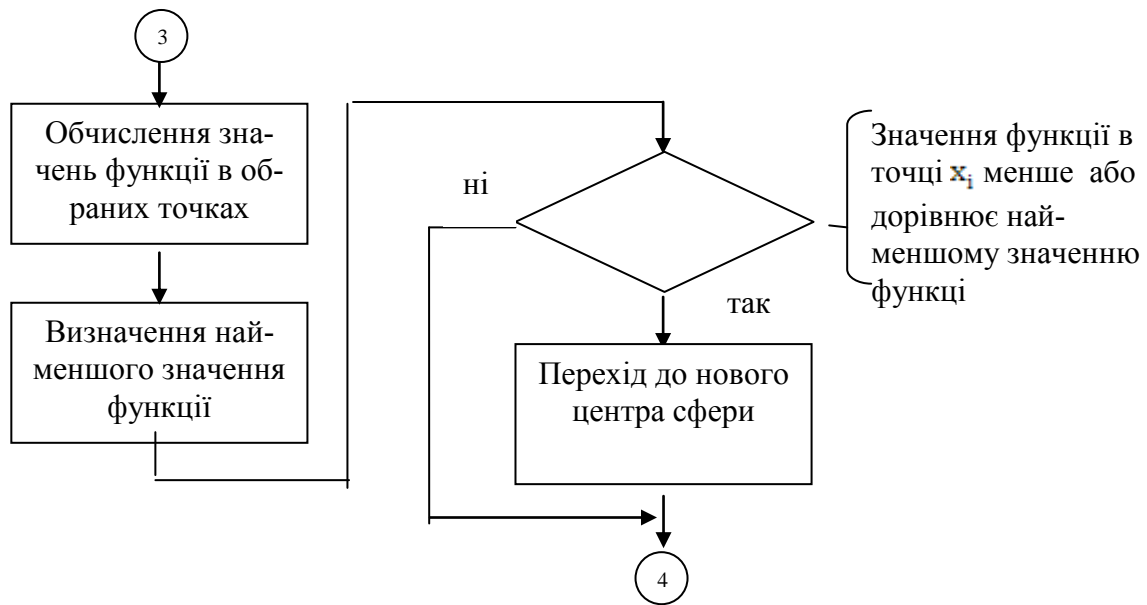


Рис. 2. Продовження рис. 1.

Перебираючи значення функцій в точках, що належать площині, знайдемо точку, в якій функція матиме найменше значення – це точка з координатами (1.2;4.6), де функція приймає значення $f(x_1, x_2)=13.4$.

Подальші дії відбуватимуться згідно алгоритму. В результаті одержимо відповідь: точка з координатами (4;3), в якій функція приймає значення $f(x_1, x_2)=3$.

Послідовність дій знаходження мінімуму функції проілюстровано на рис. 3.

Висновки

Запропонований метод пошуку оптимального значення функції являється досить простим і легким в реалізації. Але беручи до уваги те, що в основі методу лежить випадковий вибір напрямку пошуку, немає можливості вказати кількість ітерацій та швидкодію методу.

Щоб підвищити точність обчислень можна вдатися до таких дій:

- збільшити кількість точок, що вибираються на поверхні кола або сфери, що надасть можливість розглянути більше напрямків;
- збільшити кількість повторів вибору довільних точок на поверхні фігури при невдалому пошуку точки \tilde{x}_n ;

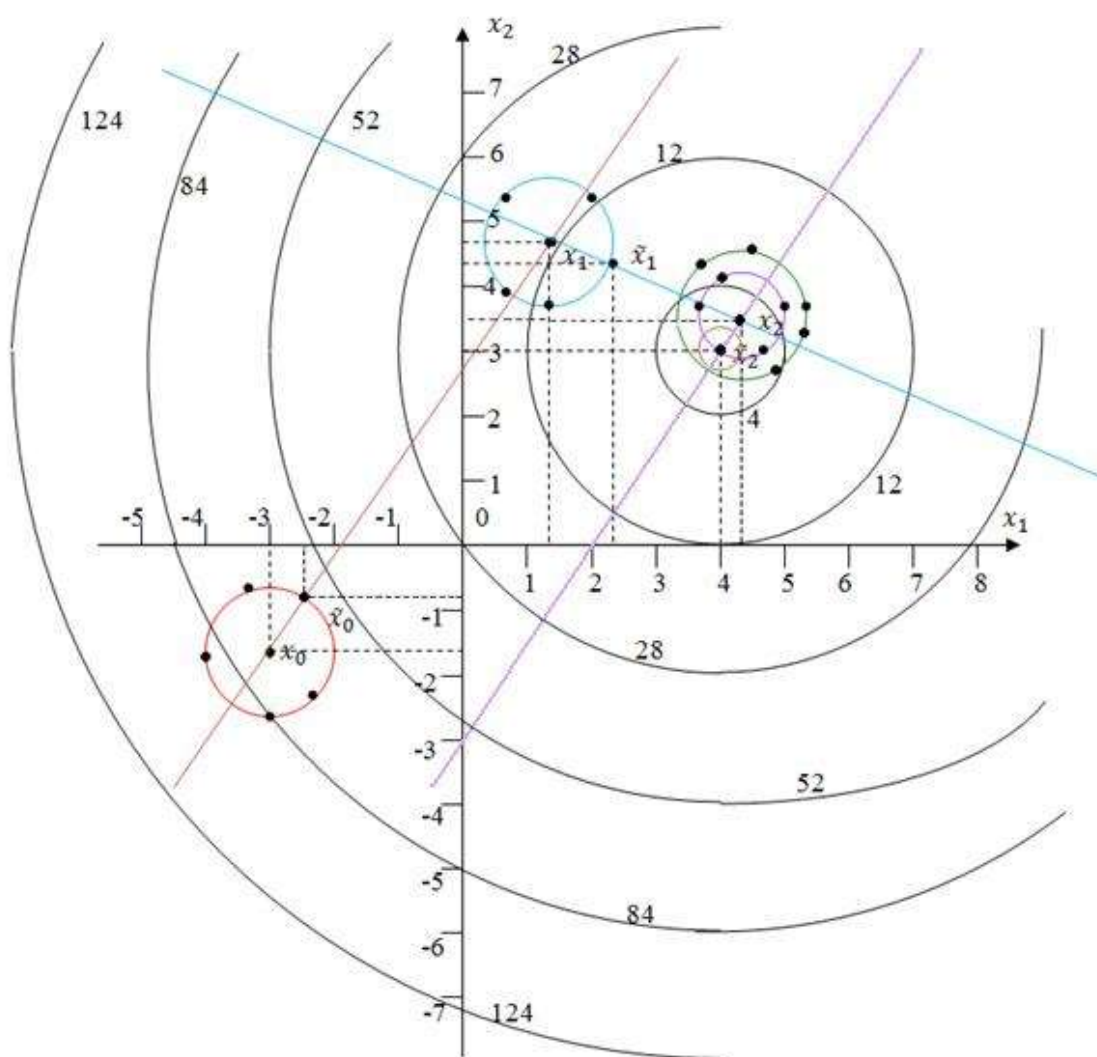


Рис. 3. Геометричне представлення пошуку мінімального значення заданої функції за допомогою запропонованого методу.

- зменшити крок перевірки α на прямій $x_n \tilde{x}_n$ – вибір точок на прямій $x_n \tilde{x}_n$ проходить з меншим кроком, що зменшує імовірність втрати «кращого» значення функції;
- зменшити величини β , на яку змінюється радіуса фігури – радіус кола або сфери, всередині якої знаходиться оптимум, буде змінюватись з меншим кроком, що дасть можливість більш точно дослідити дану область.

Всі ці дії призводять до збільшення точності обчислень, але з тим збільшується і час, необхідний для проведення пошуку оптимального значення.

Запропонований метод доцільно використовувати в області прогнозування та розрахунків, а також, в тих випадках, коли натуральний експеримент немож-

ливий або обтяжливий за тими чи іншими умовами. Він знайде застосування при розробці конструкцій виробів приладобудування.

Література

1. Химмельблау Д. Прикладное нелинейное программирование / Пер. с англ. – И.М.Быховской и В. Т. Вавилова: Мир, 1975. – С. 193-207.

<p>Левандовская И.В. Метод случайного поиска оптимума с переменным радиусом поиска</p> <p>В статье приведен способ нахождения оптимального значения функции при помощи метода случайного поиска с переменным радиусом и случайным направлением. Показано преимущества предложенного оптимизационного алгоритма определения экстремальных значений функций и пример его применения</p>	<p>Levandovska I.V. Method of random search of optimum with the variable radius of search</p> <p>In article the way of a finding of optimum value of function by means of a method of casual search with variable radius and a casual direction is resulted. It is shown advantages of the offered optimising algorithm of definition of extreme values of functions and an example of its application.</p>
--	--

*Надійшла до редакції
18 березня 2009 року*