

УДК 681.518:004.93'.1

А.С. Довбиш¹, О.Б. Берест², Р.С. Приходченко³¹СумДУ, м. Суми, Україна, kras@id.sumdu.edu.ua²СумДУ, м. Суми, Україна, Berest_Oleg@mail.ru³СумДУ, м. Суми, Україна, hiro-pr@yandex.ru

ІНФОРМАЦІЙНО-ЕКСТРЕМАЛЬНИЙ АЛГОРИТМ НАВЧАННЯ КЕРОВАНОГО ГОЛОСОМ МАНІПУЛЯТОРА

Запропоновано модель інтелектуальної руки-маніпулятора, яка здатна розрізняти голосові команди. Розпізнавання голосових команд здійснюється в рамках інформаційно-екстремальної інтелектуальної технології, яка базується на максимізації інформаційної спроможності системи в процесі її навчання. Для заданого алфавіту голосових команд сформовано вхідний математичний опис руки-маніпулятора та реалізовано інформаційно-екстремальний алгоритм її навчання. За результатами фізичного моделювання підтверджено працездатність розробленого інформаційного та програмного забезпечення.

МАНІПУЛЯТОР, ГОЛОСОВА КОМАНДА, ІНФОРМАЦІЙНО-ЕКСТРЕМАЛЬНА ІНТЕЛЕКТУАЛЬНА ТЕХНОЛОГІЯ, КЛАС РОЗПІЗНАВАННЯ, ІНФОРМАЦІЙНИЙ КРИТЕРІЙ, ФУНКЦІОНАЛЬНА ЕФЕКТИВНІСТЬ, ОЗНАКА РОЗПІЗНАВАННЯ

Вступ

Використання інтелектуальних маніпуляторів дозволяє розширити функціональні можливості робототехнічних комплексів різного призначення, серед яких важливого соціального значення набуває їх використання при протезуванні з метою підвищення точності імітації роботи кінцівок людини. Відомі принципи дії таких приладів базуються на використанні надчутливих нейросенсорів, що призводить до надзвичайно великої ціни протезів і, як наслідок, неможливості їх широкого застосування. Одним із перспективних підходів до розширення функціональних можливостей і підвищення точності й оперативності виконання команд є застосування голосового керування роботами-маніпуляторами [1-8]. Але існуючі технології голосового керування через довільні початкові умови керованого процесу характеризуються або наявністю помилкового виконання команд, або невисокою оперативністю. Як перспективний напрям підвищення функціональної ефективності керованої голосом руки-маніпулятора є застосування ідей і методів інформаційно-екстремальної інтелектуальної технології (ІЕІ-технології) аналізу і синтезу здатних навчатися систем керування, що ґрунтуються на максимізації інформаційної спроможності системи в процесі її навчання [9-11].

В статті розглядається інформаційно-екстремальний алгоритм самонавчання інтелектуальної руки-маніпулятора, здатної виконувати певні дії на основі розпізнавання.

1. Основна частина

Система керування за голосом маніпулятора складається з двох основних частин: системи розпізнавання голосових команд і виконавчого пристрою. Розглянемо формалізовану постановку задачі машинного навчання системи розпізнавання голосових команд. Нехай задано нечіткий у

загальному випадку алфавіт класів розпізнавання $\{X_m^o \mid m = \overline{1, M}\}$, який характеризує скінчену множину голосових команд, і структурований вектор параметрів функціонування системи керування $g = \langle g_1, \dots, g_\xi, \dots, g_\Xi \rangle$ з відповідними на них обмеженнями.

В процесі навчання системи керування необхідно сформулювати для заданого алфавіту $\{X_m^o\}$ навчальну матрицю типу «об'єкт-властивість» $\|y_{m,i}^{(j)}\|, i = \overline{1, N}, j = \overline{1, n}$, де N – кількість ознак розпізнавання і n – кількість реалізацій образу, і оптимізувати координати вектора g за умови забезпечення глобального максимуму усередненого за алфавітом класів розпізнавання інформаційного критерію функціональної ефективності (КФЕ) навчання системи в робочій (допустимій) області визначення його функції

$$\bar{E}^* = \frac{1}{M} \sum_{m=1}^M E_m^* \quad (1)$$

де E_m^* – максимальне значення КФЕ навчання системи розпізнавати реалізації класу X_m^o .

На етапі екзамену, тобто безпосередньо в робочому режимі маніпулятора, необхідно прийняти рішення про належність голосової команди до одного із класів алфавіту $\{X_m^o\}$.

Як відомо навчання системи керування в рамках ІЕІ-технології полягає в цілеспрямованій трансформації апріорного у загальному випадку нечіткого розбиття простору ознак в чітке розбиття класів розпізнавання. При цьому на кожному кроці навчання відбувається відновлення побудованих у радіальному базисі простору ознак оптимальних контейнерів, геометричні параметри яких визначають вирішальні правила, що застосовуються в режимі екзамену.

Основними етапами реалізації інформаційно-екстремального алгоритму машинного навчання системи розпізнавання голосових команд є:

- 1) формування вхідного математичного опису;
- 2) перетворення вхідної навчальної матриці при заданій системі контрольних допусків на ознаки розпізнавання в робочу бінарну навчальну матрицю, над якою в процесі навчання здійснюються допустимі перетворення;
- 3) визначення шляхом статистичного усереднення бінарної навчальної матриці для кожного класу розпізнавання еталонного двійкового вектора-реалізації, вершина якого визначає геометричний центр контейнера відповідного класу;
- 4) визначення для кожного класу розпізнавання із заданого алфавіту найближчого сусіда шляхом знаходження в матриці кодових відстаней найменшої кодової відстані між їх еталонними векторами;
- 5) зміна на кожному кроці навчання радіусів контейнерів класів розпізнавання за умови, що вони не перевищують міжцентрову відстань для найближчих сусідів;
- 6) обчислення на кожному кроці навчання інформаційного КФЕ (1);
- 7) пошук глобального максимуму критерію (1) в робочій області визначення його функції;
- 8) визначення оптимального параметра навчання, яке забезпечує максимальне значення критерію (1) в робочій (допустимій) області визначення його функції;
- 9) формування за результатами навчання системи розпізнавання чітких вирішальних правил, які для класифікатора з гіперсферичною формою контейнерів класів можна представити у вигляді функції належності двійкового вектора-реалізації, що розпізнається, наприклад класу X_m^o

$$\mu_m = 1 - \frac{d(x_m \oplus x_e)}{d_m^*}, \quad (2)$$

де $d(x_m \oplus x_e)$ – кодова відстань від вершини еталонного двійкового вектора-реалізації класу X_m^o і вектором x_e , що розпізнається; d_m^* – оптимальний радіус контейнера класу X_m^o , що визначено в процесі машинного навчання.

Розглянемо детальніше реалізацію кожного з цих етапів.

Формування вхідного математичного опису системи розпізнавання команд здійснювалося за методом статистичного розподілу мел-кепстральних коефіцієнтів (MFCC) [12], що дозволило перетворити вхідну фонограму на структурований вектор коефіцієнтів, який розглядається як вхідна реалізація команди. Оскільки вхідні вектори-реалізації навіть для однакової команди можуть мати різну довжину, то для їх нормалізації було застосовано метод, запропонований в праці [13].

Подальше формування навчальної матриці полягало в записі акустичних сигналів команд для заданого алфавіту класів розпізнавання. Акустичні сигнали створювалися за допомогою додатка

Audacity [14] в моно-режимі з частотою дискретизації 16000 та зберігались в WAVE форматі. Для побудови векторів-реалізацій була використана модифікована бібліотека SPro 5.0, написана на об'єктно-орієнтованій мові C++.

Навчання маніпулятора здійснювалося за інформаційно-екстремальним алгоритмом, описаним в працях [9-11, 15-16]. Як інформаційний КФЕ навчання системи розпізнавати реалізації класу X_m^o розглядалася модифікована інформаційна міра Кульбака [15]

$$E_m^{(k)} = \log_2 \left(\frac{2 - (\alpha_m^{(k)}(d) + \beta_m^{(k)}(d))}{\alpha_m^{(k)}(d) + \beta_m^{(k)}(d)} \right)^* * [1 - (\alpha_m^{(k)}(d) + \beta_m^{(k)}(d))] \quad (3)$$

де $\alpha_m^{(k)}(d)$ – помилка першого роду при прийнятті рішень на k -му кроці навчання; $\beta_m^{(k)}(d)$ – помилка другого роду; d – дистанційна міра, що визначає радіуси гіперсферичних контейнерів класів розпізнавання.

Таким чином, алгоритм навчання системи полягає в реалізації процедури знаходження інформаційного критерію в робочій області його визначення з метою побудови безпомилкових за навчальною матрицею вирішальних правил.

Функціональна схема керування здатного навчатися маніпулятора показана на рис. 1.



Рис. 1. Функціональна схема керування здатного навчатися маніпулятора

Наведена на рис. 1. структура дозволяє функціонувати маніпулятору у режимах машинного навчання і екзамену, тобто безпосереднього розпізнавання голосових команд в робочому режимі. Звуковий сигнал від оператора потрапляє в блок первинної обробки інформації (БПОІ), який його нормалізує і перетворює до вигляду, зручного для подальшого оброблення ЕОМ. З виходу БПОІ нормалізований сигнал поступає в блок формування векторів розпізнавання, де формуються бінарні вектори-реалізації образів шляхом порівняння поточних ознак розпізнавання з їх відповідними контрольними допусками. Таким чином на виході

блоку формування векторів розпізнавання формується структурований бінарний вектор-реалізація голосової команди, кожна координата якого є одномісним предикатом, що приймає значення «1», якщо значення ознаки знаходиться в полі контрольних допусків, і «0», якщо – не знаходиться. В цьому блоці при послідовному багатократному надходженні векторів-реалізацій голосових команд формується бінарна навчальна матриця, яка поступає в блок «Навчання», де за наведеним вище інформаційно-екстремальним алгоритмом відбувається оптимізація геометричних параметрів контейнерів класів розпізнавання.

З виходу блоку «Навчання» у базу знань поступають оптимальні параметри функціонування системи, які забезпечують максимальну функціональну ефективність навчання системи розпізнавання.

В режимі екзамену у блок «Екзамен» поступають значення параметрів еталонних векторів і визначені на етапі навчання оптимальні значення геометричних параметрів контейнерів класів розпізнавання, які формують вирішальні правила – функції належності (2). Прийняття рішень на екзамені про належність вектора-реалізації, що розпізнається, одному із класів із заданого алфавіту здійснювалося в рамках ІЕІ-технології за максимальним значенням функції належності

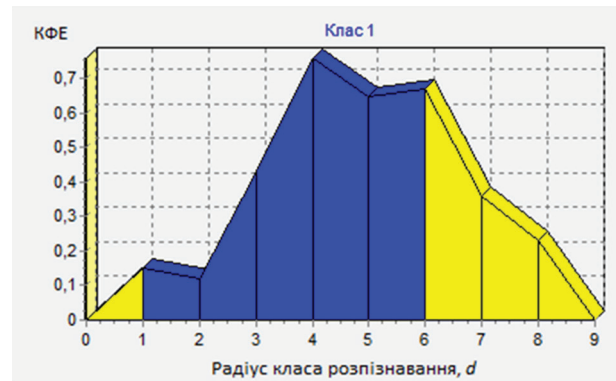
Результат екзамену передається на блок кінцевої обробки інформації (БКОІ) для формування вихідних відповідних керуючих сигналів, які поступають безпосередньо на виконавчі органи маніпулятора, визначаючи алгоритм його роботи.

Реалізація алгоритму навчання маніпулятора розглядалася на прикладі розпізнавання двох команд «ТАКЕ» та «ПУТ» Навчальна матриця складалася з 30 векторів реалізацій кожного класу, а вектори-реалізації мали по 20 ознак розпізнавання.

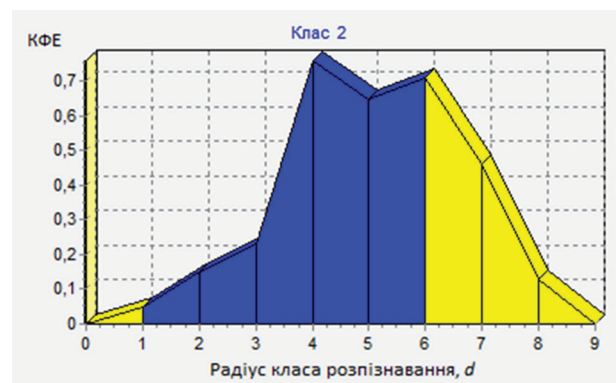
З метою аналізу інформаційно-екстремального алгоритму навчання маніпулятора розпізнавати реалізації двох класів: «ТАКЕ» – клас X_1^o і «ПУТ» – клас X_2^o досліджувалася показана на рис. 2 динаміка зміни усередненого інформаційного критерію (3) від радіусів контейнерів, що відновлювалися.

На рис. 2 темні ділянки графіків позначають робочі (допустимі) області визначення функції критерію (3), в яких значення першої та другої достовірностей більше помилок першого і другого роду відповідно. Аналіз графіків на рис. 2 показує, що при значенні параметра поля контрольних допусків на ознаки розпізнавання, яке при формуванні вхідної бінарної матриці становить $\delta_k = 45$ (у відсотках від номінального значення ознаки розпізнавання), оптимальні радіуси контейнерів дорівнюють у кодових одиницях відповідно $d_1^* = 4$ і $d_2^* = 4$. При цьому глобальний максимум інформаційного КФЕ в робочій області визначення функції (3) дорівнює $E^* = 0,72$, що свідчить про достатньо високі

асимптотичні характеристики системи керування маніпулятором, які визначаються функціональною ефективністю навчання системи.



а



б

Рис. 2. Графіки залежності усередненого КФЕ від радіусів класів: а – клас X_1^o ; б – клас X_2^o

За одержаними в процесі навчання геометричними параметрами контейнерів класів розпізнавання було побудовано вирішальні правила, які застосовувалися при виконанні рукою-маніпулятором операцій «Взяти» та «ПОКЛАСТИ». Рішення про належність реалізації, що розпізнавалася, до відповідного класу приймалося за максимальним значенням функції (2). При цьому, якщо реалізація не належала відповідному класу розпізнавання, то значення функції належності (2) для нього було від'ємним. Якщо команда була невідомою, тобто функції належності до кожного з класів були від'ємні, маніпулятор не виконував жодної дії.

Висновки

Розроблене інформаційне і програмне забезпечення інформаційно-екстремальної системи керування маніпулятором дозволило шляхом машинного навчання надати йому властивість адаптивності за умови апріорної невизначеності, обумовленої довільними початковими умовами керованого процесу та впливом неконтрольованих факторів.

Побудовані в процесі машинного навчання маніпулятора за розглянутим алгоритмом вирішальні

правила не є безпомилковими за навчальною матрицею, оскільки застосована система контрольних допусків на ознаки розпізнавання априорно є неоптимальною. З метою побудови безпомилкових за навчальною матрицею вирішальних правил згідно з принципом відкладених рішень необхідно здійснювати оптимізацію інших параметрів функціонування, які впливають на функціональну ефективність навчання СППР. Такими параметрами можуть бути, наприклад, параметри поля контрольних допусків на ознаки розпізнавання, словника ознак тощо.

Список літератури: 1. Орловський І. А. Лабораторний стенд керування маніпулятором М10П від SCADA системи TRACE MODE / І. А. Орловський, О. І. Крат, П. П. Зав'язун, Ю. С. Бірюков // Електротехніка та електроенергетика. – 2013. – № 2. – С. 54–61. 2. Конох І. С. Багатоконтрольний пристрій керування роботом-маніпулятором з електричними сервоприводами постійного струму та інкрементальними енкодерами / І. С. Конох, М. Ю. Базишин // Інженерні та освітні технології в електротехнічних і комп'ютерних системах. – 2013. – № 2. – С. 92–105. 3. Невлюдов І. Ш. Голосовое формирование управляющих команд робота в САПР технологических процессов / И. Ш. Невлюдов, А. М. Цымбал, С. С. Милютин // Восточно-Европейский журнал передовых технологий. – 2008. – Т. 2, № 2 (32). – С. 12–14. 4. Furtună F., Using Discriminant Analysis in Speech Recognition / F. Furtună, M. Dârdală // The Proceedings Of The Fourth National Conference Humman Computer Interaction Rochi 2007, Universitatea Ovidius Constanța, 2007, MatrixRom, Bucharest, 2007. 5. Divenyi P. Speech Separation by Humans and Machines. / P. Divenyi // Kluwer Academic Publishers, 2005 – 319 p. 6. Губочкин И. В. Разработка алгоритмов анализа и распознавания речи на основе адаптивной кластерной модели и критерия минимального информационного рассогласования [Текст]: автореф. дис. канд. техн. наук : 05.13.17 / И. В. Губочкин. – Нижний Новгород, 2011. – 22 с. 7. Гребнов С. В. Аналитический обзор методов распознавания речи в системах голосового управления / С. В. Гребнов // Вестник ИГЭУ – 2009. – № 3. – С. 83–85. 8. Цымбал А. М. Транслятор команд голосового управления роботом [Текст] : тезисы докл. / А. М. Цымбал, С. С. Милютин // Современные проблемы радиотехники и телекоммуникаций «РТ-2008»: 4-я Международная молодежная научно-техническая конф. – Севастополь, 2008. – С. 258. 9. Dovbysh A. S. Information-Extreme Algorithm for Optimizing Parameters of Hyperellipsoidal Containers of Recognition Classes / A. S. Dovbysh, N. N. Budnyk, V. V. Moskalenko // Journal of automation and information sciences. – New York: Begell House Inc. – v44.i10. – 2012. – P. 35–44. 10. Берест О. Б. Інформаційно-екстремальний алгоритм навчання системи керування вирощуванням сцинтиляційних монокристалів. / О. Б. Берест, А. С. Довбиш, Ю. С. Козьмін

// Тематичний випуск «Системний аналіз, управління та інформаційні технології». Харків: НТУ «ХПІ». – 2012. – № 30. – 128 с. (с. 54–60). 11. Довбиш А. С. Информационно-экстремальный алгоритм обучения системы диагностирования патологических процессов / А. С. Довбиш, С. А. С. М. Джулгам, А. А. Стадник // Сборник статей «Инновации в науке». – Новосибирск: НП «СибАК». – 2013. – № 23. – С. 45–54. 12. Sahidullah Md. Design, analysis and experimental evaluation of block based transformation in MFCC computation for speaker recognition / Md. Sahidullah; G. Saha // Speech Communication. – May 2012 – 54 (4). – pp. 543–565 13. Простой алгоритм распознавания речи по короткому словарю на основе MFCC [Электронный ресурс]. – Режим доступа: <http://habrahabr.ru/post/150251/> – 05.11.2014 г. – Загл. с экрана. 14. Audacity® is free, open source, cross-platform software for recording and editing sounds. [Electronic resource]. – Access mode: <http://audacity.sourceforge.net/?lang=ru> – 11.10.2014 г. – Screen title. 15. Довбиш А. С. Основи проектування інтелектуальних систем: навчальний посібник. / А. С. Довбиш // – Суми: Вид-во СумДУ, 2009. – 171 с. 16. Берест О. Б. Оценка функциональной эффективности обучения автоматизированной системы управления технологическим процессом / О. Б. Берест, А. С. Довбиш // Вестник СумГУ. Серия технических наук. – 2012. – № 2. – С. 39–47.

Надійшла до редколегії 11.02.2014

УДК 681.518:004.93'.1

Інформаційно-екстремальний алгоритм обучения управляемого голосом манипулятора / А. С. Довбиш, О. Б. Берест, Р. С. Приходченко // Бионика интеллекта: научн.-техн. журнал. – 2015. – № 1 (84). – С. 30–33.

Предложена модель интеллектуальной руки-манипулятора, которая способна различать голосовые команды. Распознавание звуковых инструкций происходит в рамках интеллектуальной экстремально-информационной технологии, основанной на максимизации информационной способности системы. Были сформированы обучающие матрицы и реализован базовый алгоритм обучения. В результате проведенных экспериментов удалось получить удовлетворительный показатель функционирования интеллектуальной системы.

Ил. 2. Библиогр.: 16 назв.

UDK 681.518:004.93'.1

Information-extreme learning algorithm of voice-controlled manipulator / A. S. Dovbysh, O. B. Berest, R. S. Pryhodchenko // Bionics of Intelligence: Sci. Mag. – 2015. – № 1 (84). – P. 30–33.

A model of intelligent robotic arm able to recognize the voice commands is proposed. Recognition of sound instruction is within intellectual extreme-information technology based on the maximizing of the information capacity of the system. Learning matrices were formed and basic learning algorithm was successfully implemented. As a result positive index of intellectual system functionality was obtained.

Fig. 2. Ref.: 16 items.