

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ТА КОМП'ЮТЕРНА ТЕХНІКА

УДК 681.518

М. М. Биков, канд. техн. наук, доц.; В. В. Ковтун, канд. техн. наук;
К. Конате, Ph. D.

МЕТОД ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ РОБОТИ ПАМ'ЯТІ В СИСТЕМАХ ПОШУКУ КЛЮЧОВИХ СЛІВ У МОВНОМУ СИГНАЛІ

Розроблено метод підвищення ефективності роботи структури асоціативної пам'яті для системи пошуку ключових слів в мовному сигналі на основі запропонованих принципів зберігання в окремих комірках асоціативної пам'яті тільки незбіжних частин еталонів та урахування потенційної здійсненності апріорно заданого розподілу ключових слів на класи в пам'яті еталонів. Надано математичне обґрунтування оптимального вибору асоціативної ознаки.

Вступ

Ефективність роботи системи пошуку ключових слів в мовному сигналі (KWSS) значною мірою залежить від швидкості та точності процедури порівняння реалізації з еталонами, розміщеними у пам'яті системи [1, 2]. *Метою роботи* є розробка методу підвищення ефективності роботи пам'яті KWSS, який базується на запропонованих авторами двох принципах: 1 – принцип зберігання в окремих комірках пам'яті еталонів тільки незбіжних частин інформаційних слів; 2 – принцип вибору розрядності асоціативної ознаки з урахуванням потенційної здійсненності апріорно заданого розподілу ключових слів в пам'яті еталонів на класи.

Специфіка еталонів мовних образів полягає в тому, що тільки деяка частина з них має подібні фрагменти з образом, що надійшов. Наприклад, за умови використання для опису звукотипів двійкової частотнотдетектуючої функції [3], в еталонних описах співпадають третій, четвертий і п'ятий розряди окремо для груп голосних $\{/a/, /o/, /y/\}$ і $\{/i/, /e/, /и/\}$, що дає можливість за цим фрагментом розподілити множину еталонів голосних на дві підмножини. В зв'язку із зазначеною особливістю автори пропонують зберігати еталони S мовних образів в пам'яті у вигляді масивів S^k , елементи яких подібні за визначеним двійковим фрагментом-ознакою Q_p . Це дозволяє прискорити процес розпізнавання мовного образу s_r шляхом вибірки із пам'яті за асоціативною ознакою Q_p^r масиву найвірогідніших кандидатів, а не шляхом послідовного пошуку на всій множині образів S . Вище використано такі позначення:

$$S = \{s_1, s_2, \dots, s_i, \dots, s_{m_s}\} \text{ — множина всіх еталонів,} \quad (1)$$

де

$$s_i = \{b_1^i, b_2^i, \dots, b_j^i, \dots, b_{k_b}^i\}, \quad b_j^i \in \{0, 1\} \text{ —} \quad (2)$$

i -й еталон, який можна розглядати як кортеж із k_b двійкових розрядів;

$$Q_p = \{b_{j_1}, b_{j_2}, \dots, b_{j_{m_b}}\} \text{ —} \quad (3)$$

виділена деяким чином підпоследовність розрядів цього кортежу, яка здійснює розподіл $D(S) = (S^1, S^1, \dots, S^{m_s})$ всієї множини еталонів S на m_s підмножин таким чином, що

$$\forall_{i \neq j} Q_p^i = Q_p^j \Rightarrow s_i \in S^k \ \& \ s_j \in S^k, \quad k = \overline{1, m_s}, \quad (4)$$

де Q_p – двійкова ознака – асоціативний ключ для пошуку k -ї підмножини еталонів S^k ,

$$S^k = \{s_{i_1}, s_{i_2}, \dots, s_{i_{m_k}}\}, \bigcup_{k=1}^{m_s} S^k = S. \quad (5)$$

За умови апаратної реалізації процесу класифікації за асоціативною ознакою Q_p швидкість розпізнавання значно підвищиться.

Апаратна реалізація пропонованого методу розпізнавання ключових слів вимагає розв'язання двох задач – задачі розробки структури асоціативного запам'ятовуючого пристрою (АЗП) з меншими, порівняно з відомими, апаратними затратами, а також задачі вибору асоціативної ознаки Q_p . Для розв'язання першої задачі в роботі пропонується принцип зберігання в окремих комірках пам'яті тільки незбіжних частин інформаційних слів. Асоціативна частина цих слів (ознака-ключ) зберігається в одному, спільному для всіх цих комірок, реєстрі, оснащеному логічними схемами, здатними виконувати операцію порівняння. Структура такого пристрою пам'яті і принцип його роботи описані авторами у [3], тому у цій роботі не наводяться, а розглядається тільки друга задача.

Вибір оптимальної розрядності асоціативної ознаки

Розв'язання другої задачі пов'язано з вибором числа розрядів, а також визначення їх позицій в двійковому слові. Вибір числа розрядів в асоціативній ознаці здійснювався з огляду на умову досягнення мінімуму затрат пам'яті в АЗП і вимогу отримання заданого розбиття множини еталонів на підмножини. В роботі були отримані в табличному вигляді числові значення оптимального співвідношення асоціативної і неасоціативної частин коду в еталоні.

Визначення розрядності ознаки з урахуванням тільки першої вимоги здійснюється на основі таких передумов. Нехай k_b – розрядність еталона, а m_b – розрядність ознаки Q_p . Тоді в АЗП з запропонованою структурою для зберігання повної множини еталонів, що розрізняються за неасоціативною частиною, потрібна пам'ять, що складається з O_a комірок

$$O_a = N_a \cdot 2^{N_a} + k_b - N_a, \quad (6)$$

де $N_a = k_b - m_b$ – розрядності незбіжних фрагментів еталонів.

Для розуміння прийнятих позначень приклад їх використання при розрядності двійкового еталона $k_b = 7$ показано на рис. 1. Рисочки відповідають коміркам пам'яті.

Оскільки в існуючих АЗП чи в звичайних адресних ЗУ всі двійкові слова зберігаються окремо, то для такої кількості еталонів потрібна була б пам'ять на O_p комірок:

$$O_p = k_b \cdot 2^{N_a}, \quad (7)$$

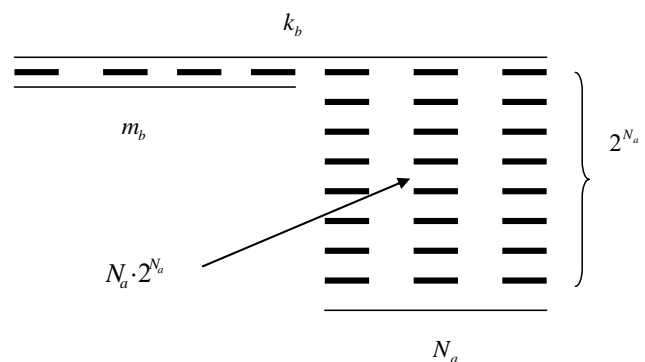


Рис. 1

звідки виграш \mathcal{E}_a в обсязі пам'яті запам'ятовуючого пристрою буде дорівнювати:

$$\mathcal{E}_a = \frac{k_b \cdot 2^{N_a}}{N_a \cdot 2^{N_a} + k_b - N_a}. \quad (8)$$

Для визначення оптимальної розрядності ознаки за критерієм (8) необхідно розв'язати відносно невідомого N_a рівняння:

$$\frac{d\mathcal{E}_a}{dN_a} = 0. \quad (9)$$

Підставляння (8) у (9) дає вираз:

$$\frac{k_b \cdot 2^{N_a} \left[\ln 2(k_b - N_a) - 1 + 2^{N_a} \right]}{(N_a \cdot 2^{N_a} + k_b - N_a)^2} = 0. \quad (10)$$

Результати розв'язання цього трансцендентного рівняння чисельним методом дихотомії наведені у вигляді графіка на рис. 2 і в табл.

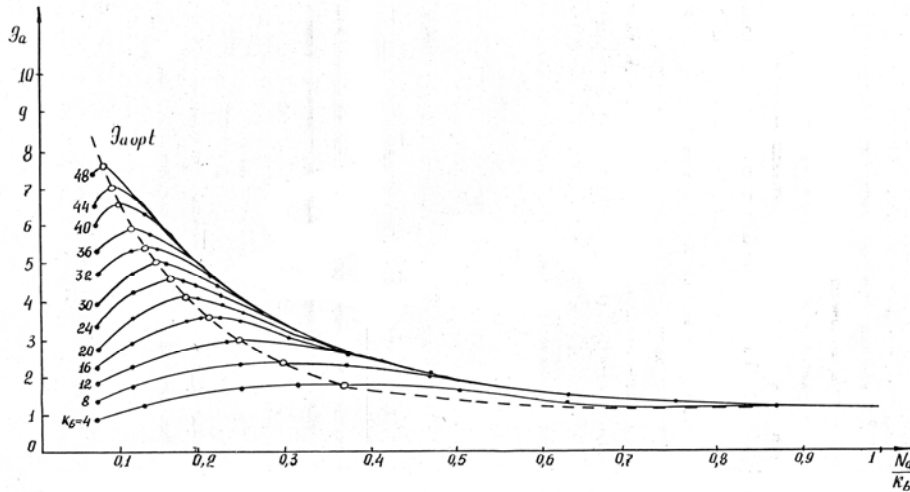


Рис. 2

В цій таблиці наведені розраховані оптимальні значення N_{aopt} і m_b розрядності інформативної частини слова і ознаки відповідно, виграш \mathcal{E}_{aopt} в обсязі пам'яті, відповідний значенню N_{aopt} , а також розрядність еталона k_b .

Таблиця дозволяє вибрати розрядність асоціативної ознаки у випадку, коли єдиним обмеженням є вимога мінімізації затрат пам'яті АЗП. З таблиці видно, що, починаючи з розрядності еталонів $k_b = 8$, потрібний в запропонованому АЗП обсяг пам'яті менший майже в два рази порівняно з відомими [4, 5] при зміні розрядності m_b асоціативної ознаки в діапазоні $[k_b - 1; 0,5 \cdot k_b]$.

В системах розпізнавання розподіл $D(S)$ множини еталонів почасти є апріорно відомим. Наприклад, на акустичному рівні системи розпізнавання мови множина можливих розподілів мовного потоку на алфавіти елементів розпізнавання задається деревом ієрархічного групування звукотипів. В цих випадках вибір розрядності асоціативної ознаки повинен здійснюватися з урахуванням потенціальної здійсненності апріорно заданого розподілу. Заданий розподіл (4) потенційно здійснимий за двійковою ознакою Q_p розрядності m_b в тому випадку, коли всю множину двійкових еталонів розрядністю k_b можна розбити не менш, ніж на m_b підмножин, потужність кожної з яких дорівнює потужності максимальної підмножини апріорного розподілу. Визначимо розрядність m_b асоціативної ознаки Q_p у відповідності до цієї вимоги. Потенціальна потужність множини двійкових векторів розрядності k_b дорівнює потужності замкнutoї множини і визначається виразом 2^{k_b} . За будь-якою підпоследовністю розрядів довжиною m_b дана множина може бути розподілена на 2^{m_b} непересічних множин, кожна з яких містить $2^{k_b - m_b}$ векторів. Тоді розрядність m_b ознаки Q_p визначиться з умови сумісного виконання двох нерівностей

Оптимальні значення розрядності асоціативної ознаки

| k_b | N_{aopt} | m_b | \mathcal{E}_{aopt} |
|-------|------------|-------|----------------------|
| 4 | 1,46 | 2 | 1,678 |
| 8 | 3,30 | 6 | 2,314 |
| 16 | 3,29 | 13 | 3,486 |
| 24 | 3,90 | 20 | 4,575 |
| 32 | 4,33 | 28 | 5,608 |
| 40 | 4,67 | 35 | 6,603 |
| 48 | 4,94 | 43 | 7,568 |
| 64 | 5,38 | 58 | 9,429 |

$$m_s \leq 2^{m_b}; \quad (11)$$

$$m_{k \max} \leq 2^{k_b - m_b}, \quad (12)$$

де $m_{k \max} = \max(m_k)$, $k = \overline{1, m_s}$.

Розв'язуючи (11) відносно m_b , а (12) відносно k_b , отримаємо математичний вираз потенціальної здійсненності апріорно заданого розподілу:

$$m_b = \lceil \log_2 m_s \rceil; \quad (13)$$

$$k_b = m_b + \lceil \log_2 m_{k \max} \rceil. \quad (14)$$

Доповнюючи вирази (13) і (14) наслідком із розв'язання рівняння (10), отримаємо таку систему обмежень:

$$m_b = \lceil \log_2 m_s \rceil;$$

$$k_b \geq m_b + \lceil \log_2 m_{k \max} \rceil; \quad (15)$$

$$k_{b \text{opt}} = \arg \min(N_a - N_{a \text{opt}}), \quad (16)$$

де вираз (16) означає, що із всіх k_b , що задовольняють обмеження (15), вибирається таке значення, за якого розрядність N_a інформативної частини відрізняється від розрядності $N_{a \text{opt}}$ на мінімальну за всіма k_b величину. Отримана система обмежень дозволяє зробити однозначний вибір розрядності асоціативної ознаки з урахуванням сумісного виконання вимоги потенціальної здійсненності апріорно заданого розподілу і вимоги досягнення максимально можливої економії пам'яті АЗП.

Висновки

Запропонований в статті метод підвищення ефективності роботи пам'яті KWSS, що ґрунтується на принципі зберігання в окремих комірках асоціативної пам'яті тільки незбіжних частин еталонів ключових слів дозволяє отримати економію в ресурсах пам'яті приблизно в 2–6 разів у разі зміни розрядності від 8 до 32 розрядів відповідно. Надане математичне обґрунтування оптимального вибору асоціативної ознаки дозволяє зробити однозначний вибір розрядності асоціативної ознаки з урахуванням сумісного виконання вимоги потенційної здійсненності апріорно заданого розподілу і вимоги досягнення максимально можливої економії пам'яті АЗП. Використання результатів цієї роботи дає змогу зручно реалізувати апаратно пристрій попередньої класифікації ключових слів в мовному потоці на флеш-мікроконтролері, в якому асоціативні ключі для пошуку можна реалізувати в пам'яті EEPROM, а еталони записати у флеш-пам'яті команд.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Comparison of Keyword Spotting Approaches for Informal Continuous Speech / Szoke I., Schwarz P., Matejka P. [and others] // In Proceedings of Interspeech 2005. — Eurospeech. Lisabon : ISCA, 2005. — P. 633—636. — ISSN 1018-4074.
2. Bykov N. M. Development of effective strategy of pattern recognition / Bykov N. M., Kuzmin I. V., Jakovenko A. I. // Proceedings of SPIE. — 2000. — Vol. 4425. — P. 76—82. — ISBN 0-8194-4807-9.
3. Биков М. Дикторонезалежне описання образів звуків в системах розпізнавання сигналів мови / Биков Микола, Раїмі Абдурахман, Биков Максим // Вісник техніки та метрології. — 2006. — Вип. 66. — С. 13—17. — ISSN 0368-6418.
4. Корнейчук В. И. Запоминающие устройства ЦВМ / Виктор Иванович Корнейчук. — Киев : Техника, 1976. — 168 с.
5. Кохонен Т. Ассоциативная память / Тейво Кохонен. — М. : Мир, 1980. — 239 с.

Рекомендована кафедрою комп'ютерних систем управління

Стаття надійшла до редакції 23.02.11

Рекомендована до друку 9.03.11

Биков Микола Максимович — професор, **Ковтун Вячеслав Васильович** — доцент.

Кафедра комп'ютерних систем управління, Вінницький національний технічний університет, Вінниця;

Конате Карім — викладач кафедри інформатики і прикладної математики.

Дакарський університет ім. Шейх Анта Діоп, м. Дакар, Сенегал