

Винахід відноситься до галузі автоматики та обчислювальної техніки і може бути використаний для автоматичного розпізнавання мовних образів.

Відомий спосіб розпізнавання образів, відповідно якому сприймають неперервний образ, що розпізнається, проводять його перетворення, виділяють характерні ознаки, виконують дискретизацію образу в послідовність елементів, проводять сегментне маркування елементів, формують двійковий опис стрічки символів, що розпізнається, різними для різних символів двійковими кодами і виконують розпізнавання за мінімумом відстані до одного з еталонних образів, для чого послідовно зчитують з пам'яті коди відстаней між кожною з порівнюваних пар символів з стрічки, що розпізнається, та еталонної стрічки, а потім додають ці відстані [Glave R.D., Vander Giet. The David speech recognition system. - Proc. IEEE Int. Conf. ASSP. - Tulsa. 1978, p. 429-432].

До недоліків даного способу розпізнавання потрібно віднести те, що він потребує додаткових витрат пам'яті на зберігання кодів відстаней, а також значного обмеження швидкості розпізнавання через втрати часу на зчитування кодів цих відстаней з пам'яті.

Найбільш близьким до способу, що заявляється, є спосіб розпізнавання образів, відповідно якому неперервний образ сприймають, перетворюють в послідовність елементів, формують двійковий опис елементів послідовності, що розпізнається, у вигляді двійкових кодів, що зберігають різниці відстаней, виконують класифікацію за мінімумом відстані до однієї з еталонних послідовностей, для чого проводять логічне порівняння послідовностей, що розпізнається, і еталонної послідовності елементів шляхом визначення відстані Хемінга між ними, на основі отриманих результатів приймають рішення про розпізнавання мовного образу [Preparata F.P., Nivergelt J. Difference-Preserving Codes. - TEE Trans. Information Theory, IT-20, 1974, p.643-649].

Недоліками даного способу є обмеження об'єму словника образів, що розпізнаються, яке обумовлене можливістю виникнення помилок через властивості кодів, що зберігають різниці, а також обмеження швидкодії через необхідність обчислення відстаней Хемінга. Наприклад, якщо елементи образу позначити символами типу «ціле», то для двійкового коду, що зберігає різниці, виконуються наступні умови:

$$1) \text{ з } |i - j| \leq t \text{ настає } H(D_i, D_j) = |i - j|,$$

$$2) \text{ з } |i - j| > t \text{ настає } H(D_i, D_j) \neq |i - j|,$$

де: $i, j \in \{1, 2, \dots, N\}$ - цілі числа (символи елементів); t - поріг, заданий цілим числом;

D_i, D_j - двійкові коди елементів i, j , які зберігають різниці; $H(D_i, D_j)$ - відстань Хемінга між двома кодovими словами.

При цьому для всякої трійки образів $\{O_{k1}, O_{k2}, O_{k3}\}$, описаної у вигляді стрічок символів:

$$O_{k1} = j_1 + 1, j_2 + 1, \dots, j_m + 1;$$

$$O_{k2} = j_1, j_2, \dots, j_m;$$

$$O_{k3} = j_1 - 1, j_2 - 1, \dots, j_m - 1,$$

Хемінгові відстані $H(O_{k1}, O_{k2})$ і $H(O_{k2}, O_{k3})$ рівні між собою. Вказана невизначеність викликає помилки в розпізнаванні при наявності в словникові образів з вказаними описами.

При виконанні класифікації за відомим способом при логічному порівнянні послідовності елементів, що розпізнається, з еталонною послідовністю необхідно визначити відстань Хемінга між ними, що знижує швидкодію вказаного способу розпізнавання образів.

Внаслідок вище зазначеного даний спосіб має невисоку надійність розпізнавання мовних образів, малий об'єм словника та низьку швидкодію.

В основу винаходу поставлена задача створення способу розпізнавання мовних образів, в якому за рахунок використання кодування образів DRP-кодами (distance rank preserving codes - кодами, що зберігають ранги відстаней) досягається збільшення надійності розпізнавання мовних образів, збільшення словника мовних образів та підвищення швидкодії розпізнавання.

Поставлена задача вирішується тим, що неперервний образ сприймають, перетворюють в послідовність елементів, формують двійковий опис елементів послідовності, що розпізнається, у вигляді двійкових кодів, що зберігають ранги відстаней між елементами, виконують класифікацію за мінімумом відстані до однієї з еталонних послідовностей, для чого проводять логічне порівняння послідовності, що розпізнається, і еталонної послідовності елементів, і на основі отриманих результатів приймають рішення про розпізнавання мовного образу.

Двійковим кодом β який зберігає ранги відстаней (DRP - кодом) є відображення $i \rightarrow \beta_i$ множини $m = \{1, 2, \dots, m\}$ в множину $\{0, 1\}^n$ двійкових послідовностей довжиною n таке, що:

$$\forall_{i, j \in m} (R(d_{ij}) = k \Rightarrow R(h_{ij}) = k),$$

де: $R(d_{ij})$ - ранг відстані d_{ij} між елементами i та j в просторі елементів;

$R(h_{ij})$ - ранг відстані h_{ij} в просторі двійкових кодів;

k - ціле число, величина рангу.

Виконання класифікації образу по запропонованому способу не потребує обчислення відстані Хемінга після логічного порівняння послідовності елементів, що розпізнається, і еталонної послідовності елементів, які описують образи. Сума одиниць в двійковій послідовності, яка визначає відстань Хемінга, може бути підрахована в відомому способі апаратним або програмним способом. При апаратній реалізації підрахунку кількості одиниць, часові затрати на визначення Хемінгової відстані будуть визначатися виразом:

$$T_x = t \cdot n \cdot m \cdot p,$$

де: t - тривалість тактового періоду;

n - розрядність двійкового DRP-коду;

m - число елементів в послідовності;

p - кількість образів в словнику образів, що розпізнається.

Виграш δ в швидкодії розпізнавання згідно запропонованого способу може бути охарактеризований скороченням часових витрат на розрахунок Хемінгових відстаней:

$$\delta = \frac{T_u}{T_n} \cdot 100\% = \frac{T_u}{T_u - T_x} \cdot 100\% = \\ = \frac{T_u}{T_u - t \cdot n \cdot m \cdot p} \cdot 100\%$$

де: T_n - час розпізнавання згідно запропонованого способу;

T_u - час розпізнавання за відомим способом.

Наприклад, для випадку розпізнавання мовних образів акустичний сигнал, який відповідає одному слову, може бути перетворений в послідовність, яка складається в середньому з 10 фонемоподібних елементів, а число образів, що розпізнаються, у словнику в середньому складає 200 слів, при цьому середня тривалість слова складає приблизно 0,8сек. Розрядність DRP-коду для кодування алфавіту А, який містить 40-60 фонемоподібних елементів, визначається співвідношенням $n \geq 7$. Якщо час розпізнавання по відомому способу T_u прийняти рівним приблизно реальному часу пред'явлення мовного образу, $T_u \approx (1 \div 1,2) T_p$, що характерно для найбільш досконалих пристроїв розпізнавання мови, а тривалість t тактового періоду прийняти рівною 2мксек, (що справедливе, наприклад, для мікропроцесорів середньої продуктивності), то виграш в швидкодії буде чисельно дорівнювати:

$$\delta = \left(1 - \frac{2 \cdot 10^{-6} \cdot 7 \cdot 10 \cdot 200}{0,8} \right) \cdot 100\% \approx 104\% .$$

Фіг.1 зображує приклад конфігурації простору елементів, що підлягають двійковому кодуванню RP-кодом.

а) - просторова конфігурація кодованої множини елементів, де цифри в колі позначають елемент, а числа над лініями визначають ранги відстаней між елементами;

б) - матриця інцидентності рангів відстаней кодованих елементів, останні позначені для зручності розрізнення римськими цифрами;

в) - DRP- код рангової конфігурації елементів

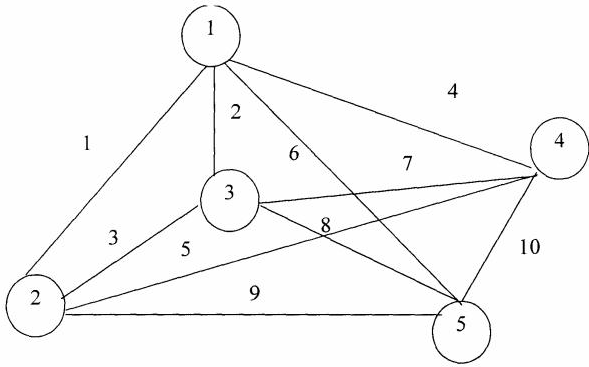
Фіг.2 - структурна схема пристрою для здійснення способу розпізнавання мовних образів.

Пристрій містить датчик сприйняття пред'явленого образу 1, який послідовно з'єднано з блоком перетворення сприйнятого образу 2, який підключено до входу блока виділення ознак 3, вихід якого з'єднують з блоком перетворення неперервного образу в послідовність елементів 4, який послідовно з'єднано з блоком формування двійкового опису 5, вихід якого підключено до входу регістра 6, який з'єднано з першим входом блоку порівняння 7, до другого входу якого підключено блок пам'яті 8, при цьому вихід блоку 8 з'єднано з блоком прийняття рішення 9.

Спосіб розпізнавання мовних образів реалізується наступним чином.

Неперервний образ сприймають, перетворюють в послідовність елементів, формують двійковий опис елементів послідовності, що розпізнається, у вигляді двійкових кодів, що зберігають ранги відстаней між елементами, виконують класифікацію за мінімумом відстані до однієї з еталонних послідовностей, для чого проводять логічне порівняння послідовності, що розпізнається, і еталонної послідовності елементів, на основі отриманих результатів приймають рішення про розпізнавання мовного образу.

Пред'явлений образ сприймається датчиком сприйняття пред'явленого образу 1, перетворюється в потрібну форму за допомогою блок перетворення сприйнятого образу 2, який з'єднується зі входом блока виділення ознак 3. На основі цих ознак в блоці перетворення неперервного образу в послідовність елементів 4 неперервний образ перетворюється в послідовність елементів, в блоці формування двійкового опису 5 формується її двійковий опис кодами, що зберігають ранги відстаней, в регістрі 6 послідовність, що розпізнається, запам'ятовується і послідовно порівнюється додаванням "по модулю 2" в блоці порівняння 7 зі всіма еталонними послідовностями, які зберігаються в блоці пам'яті 8. На основі результатів порівняння в блоці прийняття рішення 9 виконується класифікація пред'явленого образу.



a)

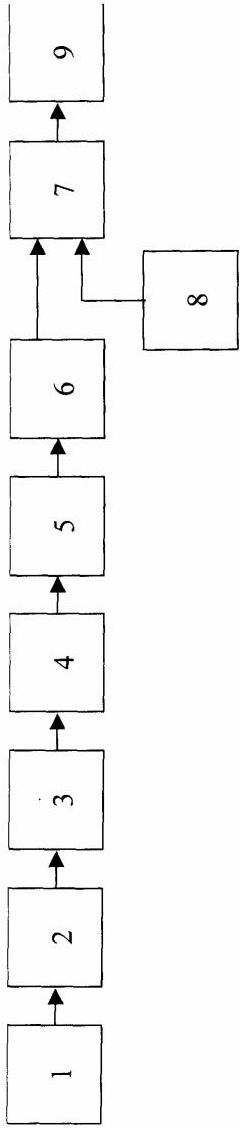
I	1	2	4	6
II	1	3	5	9
III	2	3	7	8
IV	4	5	7	10
V	6	8	9	10

б)

I	0	0	0	0	1	0	1	0	1	1
II	0	1	0	0	0	1	0	1	0	1
III	0	0	1	1	0	0	0	1	1	0
IV	1	0	0	1	0	1	1	0	0	0
V	1	1	1	0	1	0	0	0	0	0

в)

Фиг.1



Фиг.2