

УДК 004.932.751

**ЗАСТОСУВАННЯ МАТРИЧНИХ ВЕКТОРІВ ОЗНАК
У ЗАДАЧІ РОЗПІЗНАВАННЯ ДАКТИЛЬНОЇ МОВИ ЖЕСТІВ**

А.О. ГОЛІК, В.С. ДОНЧЕНКО

Представлено підхід до розв'язання задачі розпізнавання дактильної мови жестів. У повному обсязі реалізовано всі основні етапи розпізнавання: пошук та виділення на зображенні контуру долоні, пошук опуклої оболонки та «дефектів» для цього контуру, побудова векторів ознак та, власне, розпізнавання жесту. В реалізації етапу стандартизації запропоновано та реалізовано різні варіанти розв'язання, зокрема, з використанням форматів перетворень зображень. У частині формування векторів ознак запропоновано та реалізовано підхід із формуванням та використанням «матричного вектору ознак» для стандартизованого зображення контуру долоні. Представлено розвиток математичного апарату оперування з матрицями в евклідовому просторі матриць фіксованої розмірності. Таке використання матриць в якості векторів ознак є важливим, адже в багатьох прикладних задачах матриці є природними представниками об'єктів, що піддаються аналізу. Запропоновано різні варіанти встановлення відстаней відповідності для «матричних векторів ознак», зокрема, еліпсоїдальна та ортогональна.

ВСТУП

Розпізнавання об'єктів на зображенні є важливою прикладною задачею, сфера застосування якої розширюється з кожним днем. Використання спеціальних датчиків та пристроїв захоплення об'єктів на зображенні дало можливість відтворювати на комп'ютері рухи людини у реальному часі з дуже високою точністю. Досить поширеними є прикладки, які дозволяють керувати різними пристроями, наприклад проектором, за допомогою голосових команд або жестів. Головною проблемою сьогодення є той факт, що прогресивні розробки ведуться, за великим рахунком, лише у тих напрямках, які є перспективними в плані інвестування та отримання прибутку. Тому націленість виробника на широке коло споживачів є зрозумілою. У свою чергу є люди, для яких системи розпізнавання є не розвагою, а життєвою необхідністю, наприклад, глухі та глухонімі. У випадку, коли людина втратила слух у зрілому віці, зазвичай вона здатна продовжити повноцінне життя, але у значно складнішому становищі опиняються діти, які глухі з народження чи з малого віку. Розвиток таких дітей є важливою соціальною проблемою, яка, на жаль, досі не вирішена.

ПОСТАНОВКА ЗАДАЧІ

Дактильна абетка для української мови складається з 33 літер. Частина літер задається статичним жестом, інша частина включає в себе рух певного характеру. Крім того, деякі літери можуть бути показані у різних варіантах. Задача полягає у побудові системи розпізнавання, яка дозволила б розпізнавати літери абетки правильно показані різними людьми. Серед доступного обладнання є лише комп'ютер із вбудованою веб-камерою. Застосування сенсорів, різного роду датчиків та більшої кількості записуючих пристроїв не допускається у зв'язку з вимогами щодо доступності такої системи. Розпізнавання знаків дактильної мови жестів має відбуватися у реальному часі. Коли йдеться про захоплення відео потоку з веб-камери, зазвичай мається на увазі робота з упорядкованим набором зображень, адже цифровий відео потік є дискретизованим і кількість зображень за секунду чи за хвилину залежить від швидкодії системи та можливостей записуючого пристрою. Зважаючи на це, можна сказати, що розпізнавання рухомих знаків у кінцевому рахунку зводиться до розпізнавання послідовностей статичних зображень. Тому першочерговою задачею є ідентифікація жесту зі словника (абетки) на кожному окремому зображенні.

Мета роботи — формування та модельна перевірка підходів до розв'язання обговорюваної задачі, які б дозволили створення системи розпізнавання знаків дактильної мови жестів (абетки для глухонімих), яка б не потребувала додаткових пристроїв зчитування та сенсорів, була доступною для більшості людей, які її потребують, особливо дітей.

У цій статті, зокрема, наведено пропозиції, щодо варіанту побудови багатоступенчастої системи розпізнавання жестів. Власне, мова йде про реалізацію алгоритму розпізнавання на основі розширення концепції представництва об'єктів класифікації за вектором ознак: розпізнавання на базі «матричних векторів ознак».

ВИДІЛЕННЯ НА ЗОБРАЖЕННІ КОНТУРУ ДОЛОНІ

Більшість способів пошуку долоні на зображенні пов'язані з аналізом кольору пікселів. Слід зазначити, що є два основних формати представлення кольору — RGB та HSV. І якщо перший визначає колір, як комбінацію червоного, зеленого та синього, то другий — пропонує розклад на компоненти: тон, насиченість та яскравість. Хоча обидва формати мають свої недоліки та переваги, у задачах розпізнавання об'єктів на зображеннях (особливо коли йдеться про обличчя чи жести людини) використовують HSV-представлення. У загальному випадку виділення долоні за кольором шкіри складається із двох послідовних кроків. На першому — виділяють ділянки зображення, які загалом відповідають кольору шкіри людини. На другому — відбувається на вже виділених ділянках пошук тієї, яка відповідає долоні. Це не тривіальна задача, адже умови середовища — освітлення, фон тощо, мають безпосередній вплив на результати розпізнавання, тому існуючі системи вимагають наявності певного фону, чи налаштування системи під умови конкретного середовища. Згадана проблема долається використанням однокольорової рукавички (наприклад, червоної), яка дозволила б сконцентруватися безпосередньо на процесі розпізнавання. Практичну реалізацію алгоритму здійснено з використанням бібліотеки OpenCV, а саме її

оболонки для C# — EmguCV. Наявні в ній інструменти для аналізу зображення, його обробки та згладжування, пошуку контурів тощо дозволили досягти дуже високої якості пошуку контуру долоні на зображенні, в чому можна переконатися подивившись на рис. 1.



Рис. 1. Пошук та виділення на зображенні контуру долоні

КОНТУР ДОЛОНІ: ОПУКЛА ОБОЛОНКА ТА «ДЕФЕКТИ»

Інформація про виділену частину зображення відображається певним чином: у вигляді замкненого контуру, що представляє формалізований опис «захопленого» об'єкту. Контур у свою чергу зберігається у вигляді послідовності точок, кожна з яких задається певними координатами, обчисленими відносно лівого верхнього кута вихідного зображення. Зауважимо, що опис формується так, щоб зберегти інформацію, яка може бути використана для ефективного розпізнавання жестів. Зокрема, це стосується такої важливої для розпізнавання інформації, як та, що міститься в так званих «дефектах». Під дефектами зображення контуру долоні матимемо на увазі геометричні характеристики, що дозволяють ідентифікувати дактильний символ. Згадані дефекти характеризують взаємне розташування контуру долоні та його опуклої оболонки. Так, на рис. 2 дефекти представлено «криволінійними трикутниками» з синіми ділянками ламаної опуклої оболонки контуру як основами. Оскільки дефекти визначаються взаємним розташуванням контуру долоні та його опуклої оболонки, то наступним кроком обробки зображення з метою розпізнавання є побудова опуклої оболонки контуру.

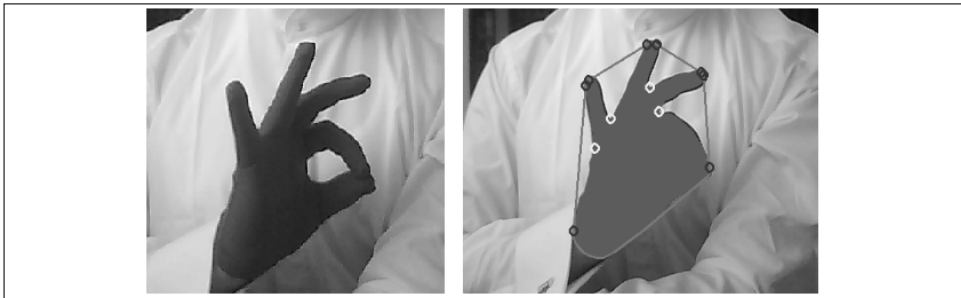


Рис. 2. Пошук та виділення контуру долоні, опуклої оболонки та дефектів

На рис. 2 проілюстровано результат роботи програми з виділення контуру долоні та побудови відповідної йому опуклої оболонки. Цей рисунок представляє чотири дефекти, для яких кружечками виділено початок, кінець та найглибша точка. Насправді рис. 2 ілюструє тільки значущі дефекти: ті, що можуть бути помітними на зображенні та, відповідно, використані для розпізнавання. Тому важливою частиною алгоритму роботи програми з виділення дефектів є сортування дефектів на «значущі» та «незначущі». Критерії значущості, якими можуть виступати, наприклад, значення довжини та

глибини дефекту, відношення довжини до глибини та деякі інші, реалізуються у спеціальних фільтрах програми.

Наостанок зауважимо, що дефекти за результатом роботи програми представлено в певному «стандартизованому» вигляді. Важлива для розпізнавання символу за дефектами інформація міститься в їхній структурі та взаємному розташуванні [1].

АЛГОРИМ РОЗПІЗНАВАННЯ ЗНАКІВ ДАКТИЛЬНОЇ МОВИ ЖЕСТІВ

Головною ідеєю, реалізованою в алгоритмі розпізнавання та, відповідно, в системі розпізнавання, є багатоетапність обробки, в ході якої здійснюється послідовне звуження множини жестів, які можуть бути розв'язком. У разі правильного налаштування системи максимальна кількість дефектів, що характеризують зображення долоні, — чотири, тому всі знаки дактильної мови жестів розбито на п'ять класів, у відповідності до цієї кількості. Класи є такими, що можуть мати спільні елементи. П'ятий клас включає в себе літери, для яких значущих дефектів не знайдено.

В цілому, алгоритм розпізнавання знаків дактильної мови жестів складається з наведених нижче етапів.

1. Віднесення жесту до одного з класів за кількістю значущих дефектів.
2. Обрахування числових характеристик для кожного дефекту в жесті.
3. Послідовне звуження множини можливих розв'язків шляхом відсіювання тих дактилів, які не відповідають встановленим інтервалам для числової характеристики, що розглядається.
4. Застосування «матричних векторів ознак» для пошуку чи перевірки остаточного результату розпізнавання.

Як вже зазначалося вище, першим необхідним елементом роботи алгоритму є виділення долоні та її контуру. Наступною складовою алгоритму є виділення та аналіз дефектів. Зокрема, аналіз дефектів включає звуження кількості можливих розв'язків, представлених зображенням, за кількістю наявних (значущих) дефектів. На наступних етапах послідовно перевіряються всі важливі властивості кожного з дефектів, що вже згадувалися вище, наприклад, структура стандартизованого представлення дефектів, довжина, глибина, відношення довжини до глибини та положення дефектів по відношенню одне до одного.

Зауважимо, що на «дефектний» опис долоні впливають певні умови фіксації зображення, зокрема кут нахилу пристрою, що записує, до долоні та положення руки. Зважаючи на це, деякі літери можуть описуватись різними наборами дефектів, та, відповідно, можуть бути віднесені одночасно до кількох із згаданих вище класів. Рис. 3 ілюструє таку ситуацію. На трьох зображеннях цього рисунку показано одну й ту саму літеру «у», але під різними кутами нахилу пристрою, який записує. Зображення *a* на рис. 3 характеризується двома дефектами, тоді як *b* та *v* — одним, але різної структури у стандартизованому представленні.

Дефекти є дуже потужним засобом аналізу, але є ситуації, коли вони вичерпують свої можливості. Серед таких випадків можна виділити наступні:

- жест не має дефектів;

- жест має дефекти, але після всіх етапів розпізнавання (пов'язаних із дефектами) однозначний розв'язок знайти не вдалося;
- є необхідність перевірити розв'язок, переконатися чи не є знайдений розв'язок хибним.

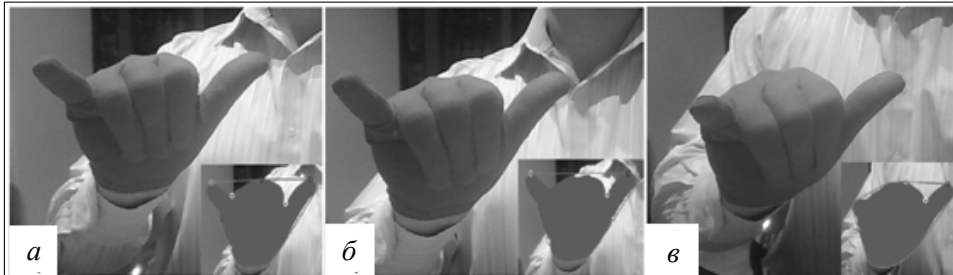


Рис. 3. Демонстрація літери «у» під різними кутами

У таких випадках актуальним є аналіз зображення контуру, зокрема побудова «матричного вектору ознак», який дозволив би порівняти зображення жестів безпосередньо. В наступних розділах мова піде саме про побудову «матричних векторів ознак» на базі зображення жесту та встановлення відстаней відповідності між ними.

СТАНДАРТИЗАЦІЯ ЗОБРАЖЕННЯ КОНТУРУ ЖЕСТУ ТА ПОБУДОВА МАТРИЧНИХ ВЕКТОРІВ ОЗНАК

Крім опуклої оболонки навколо зображення жесту може бути побудовано «мінімальний охоплюючий прямокутник» — це прямокутник мінімального розміру, в який можна вписати контур жесту. Отримавши його координати, можна вирізати із зображення ту частину, яка відповідає цьому прямокутнику та конвертувати її у бінарну матрицю. Робота з «матричними векторами ознак» вимагає використання матриць однакової розмірності. У такому випадку вони, з одного боку, є елементами лінійного простору, а з другого — представниками об'єкта, що піддається аналізу. Наприклад, можуть бути обрані квадратні матриці певного розміру. Проте охоплюючі прямокутники мають динамічний розмір, тому актуальною є проблема їх стандартизації.

Стандартизацію зручніше застосувати не до матриць, а до зображень, тому стандартизації піддається виділена прямокутна частина зображення. Стандартизація полягає в додаванні до неї прямокутних порожніх частин. Ці частини доповнюють виділений прямокутник до квадрата із стандартними розмірами. Визначення порожньої області та місце її приєднання визначається співвідношенням між висотою та шириною первісно виділеного прямокутника. Якщо висота більша ніж ширина, то додаткові порожні області потрібно додати ліворуч та праворуч, якщо навпаки — згори та знизу (рис. 4). Єдиною метою перетворення із стандартизації є збереження центрального розташування інформативно значущої частини з одночасним збільшенням зображення до потрібного розміру, тому порожні області, що додаються, однакові за геометричними розмірами. У моделі, запропонованій нижче, стандартизоване зображення є квадратом.

Саме стандартизоване зображення представляється матрицею, яку позначатимемо терміном «характеристична». Ця матриця складається з бінарних елементів: нулів чи одиниць. Одинички приписуються пікселям, що знаходяться в межах виділеного контуру долоні, нулі — поза ним. Розмір матриці визначається кількістю пікселів стандартизованого зображення за висотою та шириною. В результаті отримується квадратна бінарна («характеристична») матриця. Саме вона виступає у ролі матричного вектору ознак. Сам же підхід до розв'язання задачі розпізнавання на основі таких матричних векторів ознак можна назвати методом «характеристичних матриць» в аналізі жестів дактильної мови.

На рис. 4. наведено ілюстрацію алгоритму побудови прикладу матричного вектору ознак для дактильної літери «т».

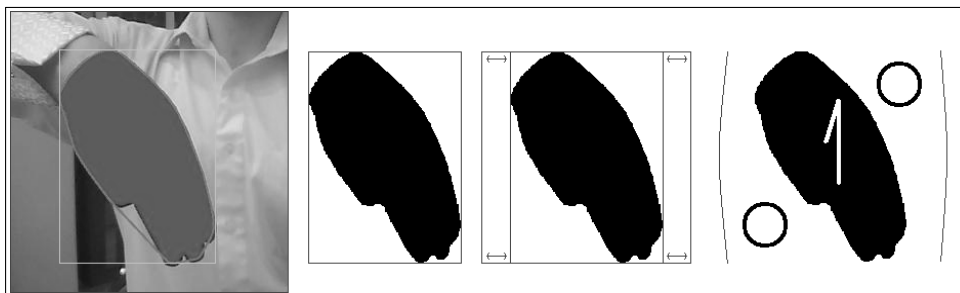


Рис. 4. Ілюстрація процесу конвертації зображення жесту в бінарну матрицю

ВІДСТАНІ ВІДПОВІДНОСТІ ДЛЯ МАТРИЧНИХ ВЕКТОРІВ ОЗНАК

Після формування векторів ознак виникає необхідність у їх порівнянні, встановленні так званої відстані відповідності. У цій статті розглянуто можливість застосування еліпсоїдальної та ортогональної відстаней відповідності, їх основною особливістю є робота не з окремими еталонами, а з наборами еталонів. Кожній дактильній літері ставиться у відповідність набір еталонів записаних у різних середовищах.

Еліпсоїдальні відстані будуються для матриць як матриць лінійних операторів між матричними евклідовими просторами засобами псевдообернення для згаданих просторів. Вони реалізуються, як і у випадку векторних евклідових просторів, через так звані «групуючі оператори» теорії псевдообернення. Такі оператори визначаються за матрицею оператора A , оператора між векторними евклідовими просторами, і визначаються співвідношенням:

$$R(A) = A^+ A^{+T}, \quad R(A^T) = (A^T)^+ (A^T)^{+T} = A^{+T} A^+.$$

Принципова роль групуючих операторів полягає в тому, що вони дають можливість побудувати «мінімальні еліпси групування»: еліпсоїди, які містять всі вектори набору $a_k, k = \overline{1, n}$ та є оптимальними в певному сенсі. Зміст оптимальності еліпсів групування полягає у тому, що його вісі утворено ортонормованим набором векторів, суми квадратів проєкцій на які є максимальними, а квадрати довжин відповідних вісей співпадають із відповідними сумами квадратів проєкцій. Точніше є справедливими наступні чотири теореми [2].

Теорема 1. Для довільного набору векторів $a_k \in R^m, k = \overline{1, n}$ розв'язком оптимізаційної задачі пошуку максимуму суми квадратів проєкцій на підпростір, породжений нормованим вектором $u \in R^m : \|u\|=1$ є вектор u_1 із сингулярності (u_1, λ_1^2) сингулярного розкладу матриці $A = (a_1 : \dots : a_n)$:

$$u_1 = \arg \min_{u \in R^m : \|u\|=1} \sum_{k=1}^r \|\text{Pr}_u a_k\|^2,$$

$$\min_{u \in R^m : \|u\|=1} \sum_{k=1}^r \|\text{Pr}_u a_k\|^2 = \lambda_1^2.$$

Теорема 2. Для довільного набору векторів $a_k \in R^m, k = \overline{1, n}$ розв'язком оптимізаційної задачі пошуку максимуму суми квадратів проєкцій на підпростір, породжений нормованим вектором $u \in R^m : \|u\|=1$ є вектор u_1 із сингулярності (u_1, λ_1^2) сингулярного розкладу матриці $A = (a_1 : \dots : a_n)$:

$$u_k = \arg \min_{u \in R^m : \|u\|=1, u \perp L(u_1, \dots, u_k)} \sum_{k=1}^r \|\text{Pr}_u a_k\|^2,$$

$$\min_{u \in R^m : \|u\|=1, u \perp L(u_1, \dots, u_k)} \sum_{k=1}^r \|\text{Pr}_u a_k\|^2 = \lambda_{k+1}^2, \quad k = \overline{1, r-1},$$

де $(u_k, \lambda_k^2), k = \overline{1, r}$, як і в попередній теоремі сингулярності сингулярного розкладу матриці, побудованої із елементів досліджуваного набору векторів.

Теорема 3. Для довільного набору векторів $a_k \in R^m, k = \overline{1, n}$

$$a_k^T R(A^T) a_k \leq r_{\max}^2 < r, \quad r_{\max}^2 = \max_{k=\overline{1, n}} a_k^T R(A^T) a_k,$$

де, як і в двох попередніх теоремах, матрицю A складено із векторів набору як із стовпчиків.

Еліпсоїд теорема 3 групує вектори набору за центральним розташуванням еліпсу групування: за еліпсом, що має центром початок координат. У практичних застосуваннях центром еліпсу вибирають середнє значення \bar{a} елементів набору:

$$\bar{a} = \frac{1}{n} \sum a_k.$$

У цьому випадку групуючий оператор будується за матрицею \tilde{A} , яка будується за центрованими середніми векторами набору $\tilde{a}_k : \tilde{a}_k = a_k - \bar{a}, k = \overline{1, n}$.

Отже, справедлива наступна теорема.

Теорема 4. Для довільного набору векторів $a_k \in R^m, k = \overline{1, n}$ виконуються такі нерівності:

$$(a_k - \bar{a})^T R(\tilde{A}^T)(a_k - \bar{a}) \leq \tilde{r}_{\max}^2 \leq r, k = \overline{1, n},$$

$$r_{\max}^2 = \max_{k=1, n} \tilde{a}_k^T R(\tilde{A}^T) \tilde{a}_k.$$

В якості набору векторів використовуються, як правило, навчальні вибірки тих чи інших класів $KL_l, l = \overline{1, L}$. В якості відстаней відповідності (точніше їхніх квадратів): функціоналів $\rho^2(x, KL_l), x \in R^m, l = \overline{1, L}$, за мінімумом значення яких здійснюється віднесення до класу, можна використовувати мінімальні еліпси групування. Це означає, що відстані відповідності визначаються співвідношеннями:

$$\rho^2(x, KL_l) = (x - \bar{a}_l)^T \frac{R(A_l^T)}{r_{1\max}^2} (x - \bar{a}_l), x \in R^m, l = \overline{1, L}.$$

Саме така еліпсоїдальна відстань використовується для характеристичних матриць.

Разом із еліпсоїдальною відстанню у статті пропонується ортогональна відстань відповідності. Вона надає можливість перенести властивості псевдообернення та SVD-розкладу на випадок матричних векторів ознак.

Позначатимемо $R^{(m \times n), K}$ — евклідов простір $m \times n$ матричних кортежів довжини K $\alpha = (A_1 \dots A_K) \in R^{(m \times n), K}$ з «природним» по компонентним скалярним добутком:

$$(\alpha, \beta) = \sum_{k=1}^K (A_k, B_k)_{tr} = \sum_{k=1}^K tr A_k^T B_k,$$

$$\alpha = (A_1 \dots A_K), \beta = (B_1 \dots B_K) \in R^{(m \times n), K}.$$

Позначимо також $\wp_\alpha : R^K \rightarrow R^{m \times n}$ [3] лінійний оператор між відповідними евклідовими просторами, що задається матричним кортежем $\alpha = (A_1 \dots A_K) \in R^{(m \times n), K}$ та визначається за допомогою матричних кортежних операцій у відповідності із співвідношенням:

$$\wp_\alpha y = \sum_{k=1}^K y_k A_k, \alpha = (A_1 \dots A_K) \in R^{(m \times n), K}, y = \begin{pmatrix} y_1 \\ \dots \\ y_K \end{pmatrix} \in R^K.$$

Теорема 5 [3]. Спряженим \wp_α^* до оператора $\wp_\alpha : R^K \rightarrow R^{m \times n}$ є лінійний оператор, що, очевидно, діє у зворотному до \wp_α напрямку: $\wp_\alpha^* : R^{m \times n} \rightarrow R^K$ та визначається співвідношенням:

$$\wp_\alpha^* X = \begin{pmatrix} tr A_1^T X \\ \dots \\ tr A_K^T X \end{pmatrix}.$$

Теорема 6 [3]. Добутком двох операторів є лінійний оператор $\wp_\alpha^* \wp_\alpha : R^K \rightarrow R^K$, який задається матрицею, що її ототожнюватимемо із самим оператором, яка визначається співвідношенням:

$$\wp_\alpha^* \wp_\alpha = \begin{pmatrix} \text{tr } A_1^T A_1, \dots, \text{tr } A_1^T A_n \\ \dots \\ \text{tr } A_n^T A_1, \dots, \text{tr } A_n^T A_n \end{pmatrix}. \quad (1)$$

Зауважимо, що матриця, визначена співвідношенням (1) є матрицею Грама елементів A_1, \dots, A_K матричного кортежу $\alpha = (A_1 : \dots : A_K)$, який задає оператор \wp_α .

Сингулярний розклад для матриці (1) є очевидним: це симетрична та невід'ємно визначена матриця. Він визначається набором сингулярностей (v_i, λ_i^2) , $i, j = \overline{1, r}$: ортонормованим набором векторів $\|v_i\|=1, v_i \perp v_j, i \neq j$; $i, j = \overline{1, r}; \lambda_1 > \lambda_2 > \dots > \lambda_r > 0$, що є власними для оператора $\wp_\alpha^* \wp_\alpha : R^K \rightarrow R^K : \wp_\alpha^* \wp_\alpha v_i = \lambda_i^2 v_i, i = \overline{1, r}$.

Визначені за сингулярностями (v_i, λ_i^2) , $i = \overline{1, r}$ матриці $U_i \in R^{m \times n} : U_i = \frac{1}{\lambda_i} \wp_\alpha v_i, i = \overline{1, r}$ є елементами набору сингулярностей (U_i, λ_i^2) , $i = \overline{1, r}$ оператора $\wp_\alpha \wp_\alpha^*$.

Розглянемо сингулярний розклад кортежного оператора. Сингулярності двох операторів: $\wp_\alpha^* \wp_\alpha, \wp_\alpha \wp_\alpha^*$, очевидним чином визначають сингулярний розклад оператора \wp_α .

Теорема 7 [3]. Сингулярний розклад кортежного оператора:

$$\wp_\alpha = \sum_{k=1}^K \lambda_k U_k v_k^T.$$

Розглянемо варіант сингулярного розкладу. Приймаючи до уваги співвідношення $U_i \in R^{m \times n} : U_i = \frac{1}{\lambda_i} \wp_\alpha v_i, i = \overline{1, r}$ та його наслідок $\lambda_i U_i = \wp_\alpha v_i, i = \overline{1, r}$, маємо

$$\wp_\alpha = \sum_{k=1}^K \lambda_k U_k v_k^T = \sum_{k=1}^K (\wp_\alpha v_k) v_k^T.$$

Зауваження загального характеру: потрібний загальний варіант теореми щодо сингулярного розкладу. Це твердження мусить стосуватися загальних евклідових просторів. Його потрібно формулювати для лінійних операторів над загальними евклідовими просторами, а вигляд має бути таким.

Теорема 8 [3]. Для довільного лінійного оператора $\wp_E : E_1 \rightarrow E_2$ над парою евклідових просторів $(E_i, (\cdot, \cdot)_i), i = 1, 2$ існує набір сингулярностей $(v_i, \lambda_i^2), (u_i, \lambda_i^2), i = \overline{1, r}, r = \text{rank } \wp_E$ операторів $\wp_E^* \wp_E, \wp_E \wp_E^*$ відповідно

зі спільним набором власних чисел λ_i^2 , $i = \overline{1, r}$, що $\wp_E x = \sum_{i=1}^r u_i \lambda(v_i, x)_1$,

$$\wp_E^* y = \sum_{i=1}^r v_i \lambda(u_i, y)_2.$$

Крім того, виконуються співвідношення:

$$u_i = \lambda_i^{-1} \wp v_i, \quad i = \overline{1, r}, \quad v_i = \lambda_i^{-1} \wp_E^* u_i, \quad i = \overline{1, r}.$$

Основні оператори теорії псевдообернення для кортежних операторів: псевдо-обернення за SVD-розкладом. У відповідності до SVD-визначення псевдообернення кортежного оператора задається [3] співвідношенням:

$$\wp_\alpha^+ = \sum_{k=1}^K \lambda^{-1} v_k (U_k, \cdot)_{tr} = \sum_{k=1}^K \lambda^{-2} v_k (\wp_\alpha v_k, \cdot)_{tr}.$$

Стандартним чином за SVD-поданням кортежного оператора визначаються ортогональні проектори базових підпросторів оператора та, відповідно, групуючі оператори.

Теорема 9. Оператори, позначені як $P(\wp_\alpha^*)$, $P(\wp_\alpha)$ та визначені співвідношеннями

$$P(\wp_\alpha^*) = \sum_{k=1}^r U_k (U_k, \cdot)_{tr}, \quad P(\wp_\alpha) = \sum_{k=1}^r v_k (v_k, \cdot) = \sum_{k=1}^r v_k v_k^T$$

є ортогональними проекторами $P_{L_{\wp_\alpha^*}}$, $P_{L_{\wp_\alpha}}$ на підпростори $L_{\wp_\alpha^*}$, L_{\wp_α} можливих значень операторів \wp_α , \wp_α^* відповідно:

$$P(\wp_\alpha^*) = P_{L_{\wp_\alpha^*}}, \quad P(\wp_\alpha) = P_{L_{\wp_\alpha}},$$

а самі ці підпростори є лінійними оболонками відповідних ортонормованих наборів:

$$L_{\wp_\alpha^*} = L(U_1, \dots, U_r), \quad L_{\wp_\alpha} = L(v_1, \dots, v_r).$$

Теорема 10. Оператори $Z(\wp_\alpha^*)$, $Z(\wp_\alpha)$, є доповненнями до тотожного оператора ортогональних проекторів — $P(\wp_\alpha^*)$, $P(\wp_\alpha)$ відповідно. Таким чином, оператори

$$Z(\wp_\alpha^*)X = X - P(\wp_\alpha^*)X, \quad Z(\wp_\alpha) = E_K - P(\wp_\alpha)$$

є ортогональними проекторами відповідно на ядра операторів, згаданих вище.

Теорема 11. Квадрат відстані $\rho^2(X, L_{\wp_\alpha})$ довільної $m \times n$ матриці X до лінійного підпростору L_{\wp_α} , що є множиною можливих значень кортежного оператора \wp_α задається формулою:

$$\rho^2(X, L_{\wp_\alpha}) = (X, Z(\wp_\alpha^*)X)_{tr} = \|X\|^2 - \sum_{k=1}^r (X, U_k)_{tr}^2.$$

Теорема 12. Квадрат відстані $\rho^2(X, L)$ довільної $m \times n$ матриці X до лінійного підпростору $L = L(A_1, \dots, A_K)$, що є лінійною оболонкою набору $m \times n$ матриць A_1, \dots, A_K задається формулою

$$\rho^2(X, L) = \rho^2(X, L_{\wp_\alpha}) = (X, Z(\wp_\alpha^*)X)_{\text{tr}} = \|X\|^2 - \sum_{k=1}^r (X, U_k)_{\text{tr}}^2$$

для кортежного оператора \wp_α , породженого набором A_1, \dots, A_K :

$$\wp_\alpha = (A_1, \dots, A_K).$$

Теорема 13. Квадрат відстані $\rho^2(X, \Gamma(a, L))$ довільної $m \times n$ матриці X до гіперплощини $\Gamma(\bar{a}, L)$: $\bar{a} = \frac{1}{K} \sum_{k=1}^K A_k$, $L = L(\tilde{A}_1, \dots, \tilde{A}_K)$, $\tilde{A}_k = A_k - \bar{a}$, $k = \overline{1, K}$, породженого набором $m \times n$ матриць A_1, \dots, A_K задається формулою

$$\rho^2(X, \Gamma(\bar{a}, L)) = (X - \bar{a}, Z(\wp_{\tilde{a}}^*)(X - \bar{a}))_{\text{tr}} = \|X - \bar{a}\|^2 - \sum_{k=1}^r (X - \bar{a}, \tilde{U}_k)_{\text{tr}}^2,$$

в якій кортежний оператор $\wp_{\tilde{a}}$ визначається співвідношенням $\wp_{\tilde{a}} = (\tilde{A}_1, \dots, \tilde{A}_K)$, а $\tilde{U}_i, i = \overline{1, r}$ ортонормований набір власних матриць оператора $\wp_{\tilde{a}}^*$.

РЕЗУЛЬТАТИ ЧИСЕЛЬНИХ ЕКСПЕРИМЕНТІВ

Було проведено низку експериментів, які дозволили визначити ефективність та робастність реалізованої системи розпізнавання знаків дактильної мови жестів. Тестувалося виділення долоні на зображенні та розпізнавання жестів на різних дистанціях та фоні за умови різного освітлення. Було знайдено універсальні параметри системи, за яких вона показує стабільний результат. Саме за умови використання такого налаштування системи, яке відповідає універсальним параметрам, і проводилися всі наведені далі експерименти.

У ході першого експерименту досліджено ефективність виділення долоні на зображенні при різному освітленні та фоні. Жест показувався, як перед собою (на фоні сорочок різного кольору), так і на фоні об'єктів у кімнаті, обличчя тощо. На рис. 5 наведено декілька тестових зображень отриманих за різних умов середовища, зокрема, збоку та напроти вікна. Проілюстровані результати роботи системи свідчать про високу якість виділення долоні на різному фоні — однотонному та не однотонному, включаючи ділянки обличчя та елементів інтер'єру кімнати.

Другий експеримент було проведено з метою перевірки можливості системи розпізнавання правильно знаходити та виділяти контур долоні на різних відстанях від записуючого пристрою. Було проведено експериментальні дослідження на відстанях 90 см, 60 см, 40 см та 20 см. На рис. 6 можна побачити ілюстрацію експерименту для літери «є».



Рис. 5. Пошук та виділення долоні у різних умовах середовища

У якості рекомендованої для роботи відстані було обрано 40–60 см для універсальних параметрів системи, хоча навіть на відстані 90 см контур долоні знаходиться достатньо точно для ефективного розпізнавання.

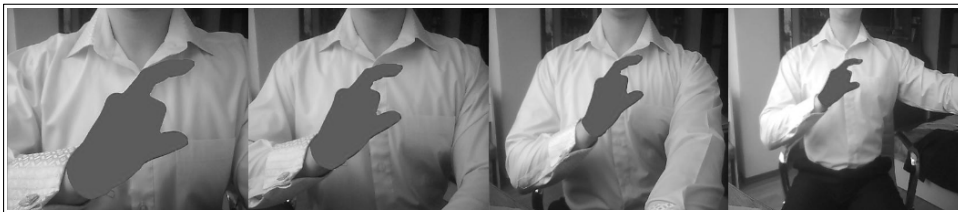


Рис. 6. Приклад жесту показаного на відстані 20 см, 40 см, 60 см та 90 см

В ході експериментів, пов'язаних із визначенням ефективності розпізнавання системи, свою доцільність продемонстрували матричні вектори ознак. Зокрема, вони дозволили ефективно вирішити проблему розпізнавання жестів, для яких кількість дефектів дорівнює нулю. Щодо роботи системи в цілому, то більшість жестів було розпізнано однозначно у всіх тестових експериментах, за виключенням випадків викликаних наявністю схожих за своєю структурою жестів, для яких кут нахилу та поворот долоні має дуже велике значення. Прикладом такого випадку є літери «о» та «р», зображені на рис. 7.



Рис. 7. Перші дві частини відповідають літері «о», другі дві — літері «р»

В залежності від положення долоні може виникнути ситуація, коли жести виглядають майже однаково за однієї відмінності — до великого пальця приєднується вказівний або середній. Розглянуті у цій статті підходи до розпізнавання не є достатньо ефективними під час розв'язання таких проблем. Серед перспективних напрямів слід виділити алгоритми скелетизації, які дозволяють отримати більш точку інформацію щодо положення окремих пальців.

Загалом, результати експериментів свідчать про високу ефективність роботи системи розпізнавання. Застосування рукавички дозволило забезпечити незалежність результатів розпізнавання від умов середовища, в тому числі фону. Природним виключенням є випадок, коли у фоні присутні області, колір яких близький до кольору рукавички.

ВИСНОВКИ

У статті запропоновано підходи до розв'язання задачі розпізнавання знаків дактильної мови жестів. Запропоновано та реалізовано два підходи до формування інформативних характеристик зображення: на основі так званих дефектів зображення жести та на основі формування матричних векторів ознак. Для другого варіанту формування представників запропоновано розвиток математичного апарату аналізу базових структур евклідових просторів для матричних евклідових просторів. Запропоновано багатоетапний алгоритм розпізнавання, який реалізовано в представленій системі розпізнавання. В роботі детально описано етапи алгоритму, що складають основу системи, в тому числі, як для випадку використання інформації щодо дефектів зображення жести, так і для матричних векторів ознак. Детально розглянуто підхід до стандартизації зображення контуру долоні та побудови матричних векторів ознак із подальшим встановленням відстаней відповідності між ними, а саме — еліпсоїдальної та ортогональної відстаней.

Проведено експериментальні дослідження, що підтвердили високу якість виділення контуру долоні на зображенні за різних умов середовища. Обидві запропоновані відстані відповідності показали свою ефективність для матричних векторів ознак. Серед знайдених недоліків — недостатня ефективність системи у випадках, коли жести літер дуже схожі між собою і для ідентифікації необхідна більш детальна інформація щодо положення окремих пальців.

Перевагою багатоетапної системи розпізнавання є той факт, що нові підходи та алгоритми можуть бути інтегровані у систему як додатковий етап та не вимагають зміни чи доопрацювання існуючих модулів. Простота у вдосконаленні та розвитку системи розпізнавання є важливим фактором в умовах актуальності цього напрямку.

ЛІТЕРАТУРА

1. *Golik A.O.* Recognition of tactile sign language: designing the feature vectors based on amount and structure of defects. XIX International Conference «Problems of Decision Making under Uncertainties». — April 23–27, 2012. — Mukachevo, Ukraine. — P. 83.
2. *Donchenko V.S.* Euclidian space of vectors and matrixes: constructive methods of description of basic structures and their usage // International Journal «Information technologies & Knowledge». — 2011. — 5, № 3. — P. 203–216.
3. *Donchenko V., Zinko T., Skotarenko F.* «Feature vectors» in grouping information problem in applied mathematics: vectors and matrixes / Problems of Computer Intellectualization. — Editors Vitali Velichko, Aleksey Voloshin, Krassimir Markov. — Kyiv, Ukrain — Sofia Bulgaria: NASU, V.M. Glushkov Institute of Cybernetics, ITHEA: 2012 — P.111–124.

Надійшла 26.14.2013