

КЛАСИФІКАЦІЯ ДАКТИЛЕМ УКРАЇНСЬКОЇ ДАКТИЛЬНОЇ АБЕТКИ

Ю.В. Крак, Г.І. Кудін, А.В. Ізотов, І.Г. Соломянюк

Київський національний університет імені Тараса Шевченка
01601, Київ, вул. Володимирська, 64/13
Yuri.krak@gmail.com

Abstract

Methods for sign language dactyloemes alphabet classification: band separation and support vector machines (SVM) are investigated. The different methods of features formation are proposed. The advantages and disadvantages of each of the proposed methods are discussed.

1. Вступ та постановка задачі

Дактильна мова використовується нечуючими людьми для передачі інформації у випадках, коли немає жестового аналога цієї інформації, зокрема - для показу імен, цифр, скорочень, географічних назв, фізичних величин тощо. Для передачі дактильної інформації використовуються рухи кисті і пальців людини. Ці рухи строго знаходяться в області видимості для співбесідника, тому при розпізнаванні дактилем потрібно аналізувати відеозображення рук людини і відслідковувати зміни їх стану або форми з плином часу [1]. В силу цього, для побудови системи класифікації дактильної інформації, в якості ознак, важливо використовувати геометричні характеристики долоні та пальців руки людини [2]. В доповіді пропонуються і досліджуються різні способи виділення і формування векторів характеристичних ознак для вирішення задач розпізнавання. Для дактилем, які важко розпізнати (це дактилеми А, В, Е, Є, І, І, Й, Л, М, Н, Р, С, Ф, Т, У, Ч, Х, Ю, Я) в силу їх подібності формування чи демонстрації, запропоновано алгоритми лінійної смугової класифікації [3] за кожним із способів виділення ознак. Також досліджується класифікація на основі дескрипторів особливих точок долоні людини з використанням гістограм орієнтованих градієнтів HOG [4] та SVM класифікатора [5].

2. Класифікація дактилем на основі алгоритму смугової роздільності

Нехай відомо до якого із класів (дактилем) відносяться вектори навчальної вибірки $x(j)$, $j = \overline{1, n}$ в просторі ознак. Запишемо умови лінійної віддільності скінченної множини точок у багатомірному просторі із умов існування розв'язку системи лінійних алгебраїчних рівнянь $x^T(j)a = y_j$, $y_j \geq \Delta$, $j = \overline{1, n}$ відносно вектора $a \in R^m$ для деяких $\Delta > 0$ і значень y_j . Для фіксованого Δ , необхідна і достатня умови існування розв'язку цієї системи, з використанням операції псевдо

обернення матриць, буде мати вигляд:

$$\min_{y \in D(\Delta)} y^T Z(X) y = y_*^T(\Delta) Z(X) y_*(\Delta) = 0,$$

$$\text{де } X = (x(1) : \dots : x(n)) = (x_{(1)} \dots x_{(m)})^T,$$

$$Z(X) = I_n - X^+ X,$$

$$D(\Delta) = \{y : y = (y_1, \dots, y_n)^T, y_j \geq \Delta, j = \overline{1, n}\}.$$

Шуканий вектор a приймає наступне значення

$$a(\Delta) = (X^T)^+ y_*(\Delta),$$

а товщина смуги δ , яка відділяє множину точок $x(j)$, $j = \overline{1, n}$ від початку координат, буде дорівнювати

$$\text{величині } y_* = \delta = \Delta / (y_*^T(\Delta) R(X) y_*(\Delta))^{1/2}, \text{ де}$$

$$R(X) = X^+ (X^T)^+.$$

Якщо покласти $\Delta = 1$, $y_*(1) = y_*$, то максимальна товщина смуги знаходиться з умови

$$y_{opt} = \arg \min_{y \in D} y^T R(X) y, \quad a_{opt} = (X^T)^+ y_{opt},$$

де $D = \{y : y^T Z(X) y = y_*^T Z(X) y_* = 0\}$ для всіх $e_j^T y \geq 1$, $j = \overline{1, n}$, $e_j^T = (\underbrace{0, \dots, 0}_{j-1}, 1, 0, \dots, 0)$.

Запишемо умови існування розв'язку і алгоритм його отримання для задачі лінійної смугової віддільності двох класів точок у просторі ознак, використовуючи введені

умови лінійної віддільності точок у R^m . Нехай відомо, що для послідовності точок

$$x(j) = (x^T_1(j) \dots x^T_m(j))^T, \quad \text{у просторі ознак}$$

$$R^m : x(j) \in R^m, \quad x_m(j) = 1, \quad j = \overline{1, n}, \text{ точки}$$

$x(i_k)$, $k = \overline{1, n_1}$ належать першому класу, а точки

$x(j_s)$, $s = \overline{1, n_2}$ - другому.

Тоді, під лінійною смуговою віддільністю цих класів розуміється існування такого вектора $a \in R^m$, для якого

$$a^T x(i_k) \geq 1, \quad k = \overline{1, n_1}, \quad a^T x(j_s) \leq -1, \quad s = \overline{1, n_2}.$$

Звідси умова лінійної смугової віддільності приймає

наступний вигляд: $\min_{y \in D} y^T Z(X)y = 0$,

$$D = \left\{ y : e_{i_k}^T y \geq 1, e_{j_s}^T y \leq -1, k = \overline{1, n_1}, s = \overline{1, n_2} \right\},$$

а значення вектора a оптимально визначається у процесі максимізації товщини розділяючої смуги

$$y_{opt} = \arg \min_{y \in D_1} y^T R(X)y,$$

де $D_1 = \left\{ y : y^T Z(X)y = 0 \right\} \cap D$ із умов

$$a_{opt} = \left(X^T \right)^+ y_{opt}.$$

Використаємо отримані умови смугової віддільності для побудови класифікатора складнороздільних дактилем.

2.1. Виділення характеристичних ознак дактилем

Для формування характеристичних ознак будемо використовувати геометро-топологічні параметри долоні руки і пальців людини при показі дактилем. Отримані параметри об'єднані у наступні шість блоків:

- 1) кути між векторами, проведеними з геометричного центру долоні до крайніх точок – зліва, справа, зверху і знизу області. Отримаємо вектор ознак розмірності 4;
- 2) нормована довжина векторів, проведених з геометричного центра до крайніх точок – вектор ознак розмірності 4;
- 3) компактність, направленість, витягнутість долоні – вектор ознак розмірності 3;
- 4) горизонтальне сканування (розбиття на 11 полос) – вектор ознак розмірності 11;
- 5) відношення кількості чорних пікселів до загальної кількості (область показу розділена на 25 клітинок) – вектор ознак розмірності 25;
- 6) дескриптори контурного аналізу – вектор ознак розмірності 4.

Для дактилем, у рамках кожного блоку, було отримано по 5 наборів. Для кожного блоку здійснено нормування векторів ознак на проміжок [0,1].

2.2. Результати класифікації для складнороздільних дактилем

Для аналізу ефективності запропонованих векторів ознак була здійснена попарна класифікація (кожна дактилема з іншими) дактилем у рамках кожного блоку ознак. Для кількісного порівняння пропонується для кожної пари дактилем визначати відносну величину через відношення між шириною полоси і сумою відстаней до відповідних сторін полоси. Експериментально встановлено, що для нормальної класифікації дане відношення має бути більше за 0,3. У Табл. 1. приведені результати попарного розділення деяких дактилем, використовуючи ознаки з різних блоків 1)-6).

Як впливає із результатів, наведених у Табл.1., що у рамках запропонованого методу класифікації і виділення геометро-топологічних ознак дактилем, самими ефективними були ознаки, що об'єднані у 1), 2), 3) і 6) блоки. Ознаки 5) і 6) блоків виявилися менш ефективними.

Таблиця 1. Класифікація складнороздільних дактилем

№ блоку ознак	Пари дактилем	Відношення	Радіуси кластерів дактилем	
2),3),6)	Всі 19	>0.30		
1)	Б,Я	0.29	0.1	1.0
	Б,Е	0.29	0.1	1.0
	Ю,К	0.26	1.0	1.5
5)	А,В	0.11	1.7	2.7
	А,Г	0.01	1.7	1.8
	В,Ю	0.11	2.7	2.3
4)	А,Б	0.24	0.66	0.9
	Я,Ц	0.06	1.1	1.6
	Я,М	0.03	1.1	1.2

3. Класифікація на основі дескрипторів особливих точок долоні людини

Розглянемо інший підхід до класифікації дактилем, а саме: шляхом отримання характеристичних ознак на основі знаходження гістограм напрямків градієнта в особливих точках зображення [4]. Процес виділення ознак і класифікації дактилем розіб'ємо на декілька етапів.

На етапі попередньої обробки виконується приведення розміру зображення до наперед заданого та перетворення його кольорового простору. Також, з метою зменшення негативного впливу шумових збурень на подальших етапах обробки, виконується розмиття зображення за Гаусом та застосування медіанного фільтра.

Другий етап – локалізація руки на зображенні. Цей етап складається з автоматичного пошуку області дактилювання, виділенні у цій області зв'язних компонент за допомогою сегментації за кольором шкіри та визначенні тієї компоненти, яка відповідає руці користувача. Після цього, для визначеної зв'язної компоненти виконується побудова мінімального обмежуючого прямокутника, який використовується для виділення на зображенні області, яка містить руку.

На третьому і четвертому етапах виконується знаходження вектора ознак з отриманого фрагмента зображення, та, відповідно, його класифікація за допомогою набору попередньо створених бінарних класифікаторів.

3.1. Підготовка даних і навчання класифікатора

Відзначимо, що із множини дактилем були взяті статичні дактилеми (тобто ті, у яких навіть при русі кисті в просторі конфігурація пальців зберігається) – це обмеження накладається вибраним зв'язком – дескриптор-класифікатор. Таким чином, до класифікації не увійшли, зокрема, дактилеми «д», «з», «ї», «ь» та ін. і класифікатор налаштований на розпізнавання 27 статичних дактилем.

Навчальна вибірка складається з 240-280 фотографій фіксованого розміру (64 на 64) для кожного жесту. Ці фотографії очищені від шумів і покривають різні масштаби і варіації зовнішнього вигляду (повороти на 2-3 градуси, нахили і т.п.). Для кожної фотографії будується вектор ознак (дескриптор) з використанням методу HOG. Далі, створюється 27 бінарних класифікаторів (лінійних SVM, 300 ітерацій, точність 0.01, завершення навчання або за кількістю ітерацій, або за досягнутою точністю) і які навчаються на отриманих дескрипторах у режимі «один проти всіх». За позитивні зразки береться вся навчальна вибірка для даного жесту.

В якості негативних – десятикратний об'єм від позитивних. Для цього із кожної із 26 вибірок, що залишилися довільним чином вибирається фіксована кількість зразків). Відзначимо, що цей попередній етап проробляється один раз, отримані дані зберігаються і в подальшому використовуються при кожному виконанні програми.

3.2. Ідентифікація дактилем

Отриманий дескриптор подається на вхід кожному класифікатору. У результаті їх роботи формується бінарний вектор розмірності 27 (по кількості дактилем, які розпізнаються) з відповідями. Якщо елемент вектора дорівнює одиниці, то це означає, що дескриптор був розпізнаний як позитивний відповідним класифікатором. На виході маємо розпізнану дактилему.

На рис. 1. показано роботу експериментальної системи при розпізнаванні дактилеми «У».

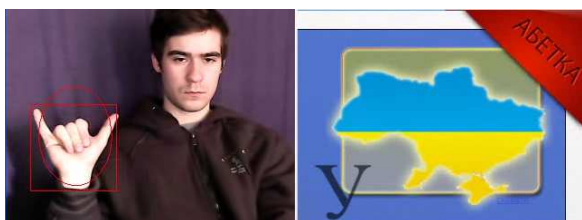


Рисунок 1. Експериментальна система класифікації дактилем української дактильної абетки

3.3. Аналіз результатів

Якісні характеристики побудованих класифікаторів досліджувалась на тестовій вибірці, яка складалась з 400 зображень для кожної дактилеми.

Для кожного з класифікаторів була оцінена точність (PR) та повнота визначення (RE), за формулами:

$PR = TP / (TP + FP)$,
 $RE = TP / (TP + FN)$, де TP – ті тестові вектори, що належали цільовому класу та були коректно розпізнані відповідним класифікатором, FP – вектори, що належали протилежному класу але були розпізнані як цільові, FN – вектори, що належали цільовому класу, але були класифіковані як ті, що належать протилежному.

Для розглянутої множини дактилем дані про точність і повноту наведені в Табл. 2.

Таблиця 2. Точність і повнота класифікації дактилем

Дактилема	Точність	Повнота
А	0.754902	0.485804
Б	0.821101	0.449749
В	0.689655	0.573066
Г	0.5	0.570866
Е	0.862245	0.330724
Ж	0.844203	0.758958
И	0.88785	0.473815
І	0.651007	0.624464
К	0.641975	0.748201
Л	0.942857	0.255814
М	0.761229	0.726862
Н	0.780731	0.795262
О	0.867876	0.863402
П	0.755396	0.253012
Р	0.757764	0.29902
С	0.696471	0.925
Т	0.527473	0.120907
У	0.789598	0.869792
Ф	0.623472	0.826924
Х	0.731928	0.883636
Ц	0.261438	0.678421
Ч	0.606516	0.94902
Ш	0.420366	0.823231
Щ	0.531617	0.610215
Ю	0.738764	0.630695
Я	0.710472	0.690619

4. Висновки

В роботі, для класифікації елементів дактильної абетки жестової мови запропоновані різні способи виділення характеристичних ознак. Досліджено два методи класифікації: метод смугової роздільності і на основі лінійних SVM, показано їх можливості для вирішення поставленої проблеми. У подальших дослідженнях буде збільшена навчальна вибірка для вибору ознак і використані оптимізаційні процедури для покращення процесу класифікації.

5. Література

- [1] Sing Languages. Ed. by D. Brentari. Cambridge: Cambridge University Press, 2010, – 714 p.
- [2] Крак Ю.В., Шкільнюк Д.В. Аналіз елементів дактильної жестової мови // Штучний інтелект. – 2010. – №3. – С.322-328.
- [3] Кириченко Н.Ф., Кудин Г.И. Анализ и синтез систем классификации сигналов средствами возмущений псевдообратных и проекционных операций // Кибернетика и системный анализ. – 2009. – №3. – С.47-57.
- [4] Dalal N., Triggs B. Histograms of oriented gradients for human detection // Computer Vision and Pattern Recognition. IEEE Computer Society Conference. – 2005. – № 1. – P. 886-893.
- [5] Hsu Chih-Wei. A Practical Guide to Support Vector Classification / Hsu Chih-Wei, Chang Chih-Chung, Lin Chih-Jen // San Francisco: Elsevier, 2008. – 216 p.