

ЗАСТОСУВАННЯ КВАДРАТА ВІДСТАНІ МАХАЛАНОВІСА В ЗАДАЧІ РОЗПІЗНАВАННЯ ОБЛИЧ

Трухов А.С.¹; Приходько С.Б., д.т.н, проф.²

^{1,2} Національний університет кораблебудування імені адмірала Макарова

^{1,2} Україна, Миколаїв

¹ timka357@gmail.com; ² sergiy.prykhodko@nuos.edu.ua

Анотація. Розроблено математичну модель для розпізнавання облич з використанням квадрата відстані Махалановіса для нормалізованих даних за допомогою перетворення Бокса-Кокса.

Ключові слова: розпізнавання образів; розпізнавання облич; математичні моделі; перетворення Бокса-Кокса; відстань Махалановіса.

Вступна частина. Методи розпізнавання образів — область технічних наук, пов'язана з розробкою принципів і побудовою систем, призначених для визначення приналежності об'єкта до певного класу. Досліджуваний об'єкт описується набором визначальних характеристик. Вибране правило прийняття рішення має ґрунтуватися на використанні обмеженої кількості характеристик, які є найбільш важливими для відмінності одного об'єкта від іншого [1]. До задач розпізнавання образів відноситься розпізнавання облич. Для вирішення цієї задачі існує багато методів, в тому числі й на основі квадрата відстані Махалановіса [2, 3]. Як зазначає ряд авторів, застосування квадрата відстані Махалановіса обумовлено нормальністю багатовимірного закону розподілу [3, 4].

Актуальність теми. Розпізнавання облич є одним із найактуальніших напрямів автоматизованої обробки інформації. Стрімкий розвиток технічного прогресу зумовив використання комп'ютерних систем та інформаційних технологій для розпізнавання облич у найрізноманітніших сферах нашого життя.

Метою роботи є побудова математичної моделі для розпізнавання облич на основі квадрата відстані Махалановіса для нормалізованих даних.

Основна частина. Кожен образ як представник відповідного класу задається за допомогою певного вектору характеристик $x = (x_1, x_2, \dots, x_n)^T$ заданим у просторі ознак X . Для вирішення задач розпізнавання образів широко застосовується класифікація на основі відстані Махалановіса [2], що обрховується за формулою (1).

$$D(x_i) = \sqrt{(x_i - \bar{x})^T S^{-1} (x_i - \bar{x})}, \quad (1)$$

де $\bar{x}$ - вибіркве середнє;

S - вибіркова коваріаційна матриця даних, яка обчислюється за формулою (2).

$$S = M[(x - M(x))(x - M(x))^T]. \quad (2)$$

Відстань Махалановіса приблизно відповідає розподілу хі-квадрат та враховує кореляцію між змінними, тобто, враховується розкид даних відносно центру мас, що має форму еліпсоїда, який описує простір допустимих значень, тому усі спостереження одного класу повинні лежати в межах цього еліпсоїду [3], в такому випадку правило прийняття рішення набуває вигляду:

$$D^2(x_i) \leq X^2(q, k), \quad (3)$$

де q – рівень значущості;

k – кількість ступенів свободи.

Формування вектору характеристик. Кожна реалізація задається вектором ознак, кожен елемент якого – нормована характеристика. В якості характеристик виступають відстані між ключовими точками обличчя. Ці точки мають бути максимально інваріантними до невеликих змін. Характеристики обираються таким чином, щоб при зміні положення, емоцій, освітлення та інших зовнішніх умов мінімально спотворювати значення вектору ознак, при цьому симетричні відстані усереднюються [5]. Для пошуку обличчя на зображенні, а також виділення ключових точок використовується ряд емпіричних алгоритмів, що виконують локалізацію особи на зображенні і потім окремих її ділянок.

За допомогою програми мовою Python та бібліотекою комп'ютерного зору Dlib для кожної фотографії набору даних [6] було сформовано вектор ознак. Dlib — система з відкритим вихідним кодом, що дозволяє безкоштовно використовувати її в будь-якій розробці. Для успішної ідентифікації обличчя може бути повернуто більш ніж на 30 градусів. Контур ока описується шістьма точками, а загальна карта обличчя описується 68 точками, що дозволяє точно порівнювати та ідентифікувати обличчя [7].

В роботі [8] автор пропонує використовувати 15 характеристик для кожного обличчя. Недоліком бібліотеки Dlib є те, що вона погано розпізнає краї обличчя, через що в залежності від повороту голови відбувається значне спотворення вектору характеристик, тому на основі цієї роботи було відібрано 11 найбільш інформативних ознак.

Задача розпізнавання значно ускладнюється зі збільшенням ознак, тому було вирішено провести розпізнавання на основі 6 характеристик: x_1 – усереднена відстань від очей до середини носа, x_2 – усереднена відстань від очей до центру рота, x_3 – усереднена відстань від очей до центру брів, x_4 – усереднена відстань від брів до верху носа, x_5 – усереднена відстань від куточків очей до верху носа, x_6 – відстань між очима. Для мінімізації впливу положення обличчя на зображенні та його віддаленості використовується нормування за допомогою ділення характеристики на базову відстань між очима [9], тому кінцевий вектор ознак має вигляд $[x_1/x_6, x_2/x_6, x_3/x_6, x_4/x_6, x_5/x_6]$.

Розпізнавання облич. Для розпізнавання було відібрано набір даних, що містить 104 фотографії однієї людини та 99 іншої.

За допомогою (3) виконується класифікація реалізацій до класу 1 або 2. При цьому кількість помилок першого роду становить 2,9%, а другого – 59,5%. Така кількість помилок розпізнавання обумовлена тим, що використання квадрата відстані Махаланобіса базується на припущенні про нормальність розподілу даних. У сфері статистики припущення про нормальність є важливим, оскільки багато статистичних методів виконують обчислення, припускаючи, що дані розподілені нормально. Обраний набір даних є ненормально розподіленими, тому виникає необхідність використання нормалізуючих перетворень.

Для підвищення ймовірності розпізнавання на основі перетворення Бокса-Кокса [10] було отримано випадкові величини, розподіл яких наближено є нормальним.

Користуючись правилом прийняття рішення (3), виконується класифікація нормалізованих реалізацій. При цьому кількість помилок першого роду становить 1,9%, а другого – 33,3%. Таким чином нормалізуюче перетворення покращило математичну модель, знизивши відсоток помилок першого роду з 2,9% до 1,9%, а другого роду з 59,5% до 33,3%.

Висновки. В ході роботи було розроблено математичну модель для розпізнавання облич, яка базується на відстані Махаланобіса. Досліджено, що математична модель, яка побудована для нормалізованих даних має більшу вірогідність правильного розпізнавання

облич. В подальшому для підвищення точності розпізнавання планується підвищити кількість ознак до 11.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- [1] Garain, U., Chaudhuri, B.B., & Ghosh, R.P. (2004). A multiple-classifier system for recognition of printed mathematical symbols. In *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004*. (Vol. 1, pp. 380-383). IEEE.
- [2] Fraser, A.M., Hengartner, N.W., Vixie, K.R., & Wohlberg, B.E. (2003). Incorporating invariants in Mahalanobis distance based classifiers: Application to Face Recognition. In *Proceedings of the International Joint Conference on Neural Networks, 2003*. (Vol. 4, pp. 3118-3123). IEEE.
- [3] Van Aelst, S., & Rousseeuw, P. (2009). Minimum volume ellipsoid. *Wiley Interdisciplinary Reviews: Computational Statistics*, 1(1), 71-82.
- [4] Gallego, G., Cuevas, C., Mohedano, R., & Garcia, N. (2013). On the Mahalanobis distance classification criterion for multidimensional normal distributions. *IEEE Transactions on Signal Processing*, 61(17), 4387-4396.
- [5] Starovoitov, V., Samal, D., Votsis, G., & Kollias, S. (1999). Face recognition by geometric features. In *Proceedings of 5-th Pattern Recognition and Information Analysis Conference*.
- [6] Burak. (2019). Pins Face Recognition. <https://www.kaggle.com/datasets/hereisburak/pins-face-recognition>
- [7] Boyko, N., Basystiuk, O., & Shakhovska, N. (2018). Performance evaluation and comparison of software for face recognition, based on dlib and opencv library. In *2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP)* (pp. 478-482). IEEE.
- [8] Starovoitov, V.V., Samal, D.I., & Briliuk, D.V. (2002). Three approaches for face recognition. In *The 6-th International Conference on Pattern Recognition and Image Analysis October* (pp. 21-26).
- [9] Starovoitov, V.V., & Samal, D. (1999). A geometric approach to face recognition. In *NSIP* (pp. 210-213).
- [10] Sakia, R.M. (1992). The Box-Cox transformation technique: a review. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 41(2), 169-178.

Trukhov A.S., Prykhodko S.B.

Application of the square of the mahalanobis distance in the problem of face recognition

Abstract. A mathematical model was developed for face recognition using the squared Mahalanobis distance for normalized data using the Box-Cox transformation.

Keywords: pattern recognition; face recognition; mathematical models; Box-Cox transformation; Mahalanobis distance.