



НОВІ ЗАСОБИ КІБЕРНЕТИКИ, ІНФОРМАТИКИ, ОБЧИСЛЮВАЛЬНОЇ ТЕХНІКИ ТА СИСТЕМНОГО АНАЛІЗУ

УДК 004.93:004.89

В.М. ОПАНАСЕНКО

Інститут кібернетики ім. В.М. Глушкова НАН України, Київ, Україна,
e-mail: vlopanas@ukr.net.

Ш.Х. ФАЗИЛОВ

Науково-дослідний інститут розвитку цифрових технологій та штучного інтелекту,
Ташкент, Узбекистан,
e-mail: sh.fazilov@gmail.com.

С.С. РАДЖАБОВ

Інститут фундаментальних та прикладних досліджень при Національному дослідному
університеті «ТШМСГ», Ташкент, Узбекистан,
e-mail: s_radjabov@yahoo.com.

Ш.С. КАХАРОВ

Кокандський університет, Коканд, Узбекистан,
e-mail: sh.kaxarov93@gmail.com.

БАГАТОРІВНЕВА СИСТЕМА РОЗПІЗНАВАННЯ ОБЛИЧЧЯ¹

Анотація. Розглянуто проблему біометричної ідентифікації особи на основі покомпонентного розпізнавання обличчя. Показано, що систему розпізнавання обличчя можна уявити як ієрархічно організовану багаторівневу систему, в якій ансамбль локальних класифікаторів формує «м'які» рішення про належність зображень окремих компонентів особи заданим класам, а потім на основі інтеграції цих рішень здійснюється формування остаточного рішення про належність розпізнавання обличчя одному із заданих класів. Сформульовано та розв'язано задачу побудови моделі локальних класифікаторів, а також вибору інтегратора проміжних розв'язків локальних класифікаторів.

Ключові слова: розпізнавання образів, багаторівнева система розпізнавання, ансамбль класифікаторів, правило комбінування класифікаторів, прийняття рішень.

ВСТУП

Останніми роками інтенсивно розроблюється один із перспективних напрямів комп'ютерного зору — розпізнавання облич. Серед наукових досліджень та розроблень, проведених у рамках цього напрямку, слід виокремити [1–6].

У практично важливих програмах розпізнавання осіб, наприклад застосунках із забезпечення безпеки, неправильна класифікація може спричинити негативні наслідки. Одним із підходів до підвищення надійності таких систем є використання ансамблю алгоритмів розпізнавання з подальшим поєднанням їхніх результатів на основі будь-якого правила інтеграції.

Як основу реалізації зазначеного підходу можна використовувати концепцію багаторівневого розпізнавання об'єкта, який має складну ієрархічну структуру. Такий об'єкт є сукупністю елементів різних рівнів ієрархії. Об'єднання окремих елементів складається з набору елементів попереднього рівня, який своєю чергою є невід'ємною частиною об'єднання елементів вищого рівня.

¹ Роботу виконано в рамках проєкту ФЗ-20200929308 Міністерства інноваційного розвитку Республіки Узбекистан.

© В.М. Опанасенко, Ш.Х. Фазілов, С.С. Раджабов, Ш.С. Кахаров, 2024

Враховуючи, що обличчя людини можна охарактеризувати як складовий об'єкт, представлений такими основними компонентами, як брови, очі, ніс, рот, праве око та ін., систему розпізнавання облич можна представити як багаторівневу систему.

ПОСТАНОВКА ЗАДАЧІ

На рис. 1 наведено спрощену схему дворівневої системи покомпонентного розпізнавання облич, яка складається з таких елементів: B_j — блок виокремлення j -го компонента особи та формування вектора ознак $X_j = (x_1^{(j)}, \dots, x_{n_j}^{(j)})$, C_j — локальний класифікатор «м'якого» розпізнавання j -го компонента, що формує оцінки близькості $Y_j = (y_1^{(j)}, \dots, y_{n_j}^{(j)})$ цього компонента до кожного із заданих класів; D — блок інтеграції локальних розв'язків, що формує розв'язок Z за обличчям, що розпізнається.

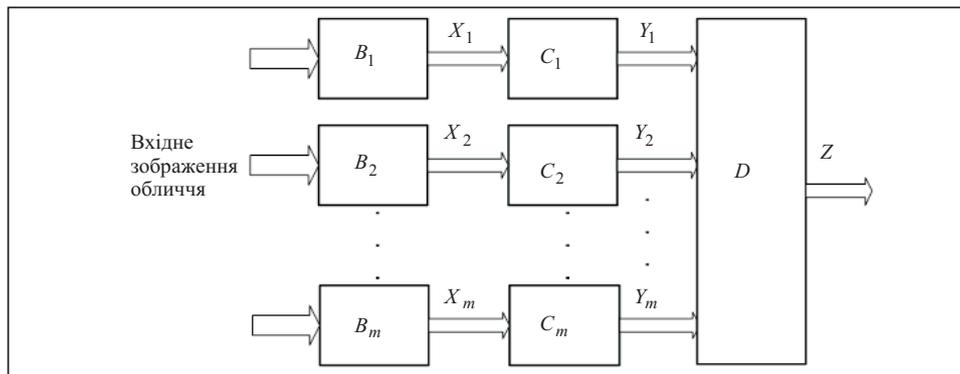


Рис. 1. Схема дворівневої системи покомпонентного розпізнавання обличчя

З наведеної схеми (див. рис.1) видно, що створення багаторівневої системи розпізнавання облич полягає у побудові локальних класифікаторів, а також інтегратора на основі вибраного правила комбінування цих класифікаторів.

Побудова локального класифікатора, який навчається з учителем, є класичною задачею розпізнавання, яку можна сформулювати в такий спосіб [7].

Нехай задано множину \mathcal{Y} допустимих об'єктів. Кожен об'єкт $S_i \in \mathcal{Y}$ описаний у ознаковому просторі великої розмірності: $S_i = a_{i1}, a_{i2}, \dots, a_{in}$.

Множина \mathcal{Y} розбита на l класів, що не перетинаються, K_1, K_2, \dots, K_l . При цьому розбиття \mathcal{Y} визначено не в повному обсязі, а є лише деякою початковою інформацією I_0 про класи K_1, K_2, \dots, K_l (навчальна вибірка):

$$I_0 = \{S_1, \dots, S_i, \dots, S_m; \tilde{\alpha}(S_1), \dots, \tilde{\alpha}(S_i), \dots, \tilde{\alpha}(S_m)\},$$

$$\tilde{\alpha}(S_i) = (\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{in}),$$

де S_i — об'єкт навчальної вибірки ($i = \overline{1, m}$); $\tilde{\alpha}(S_i)$ — інформаційний вектор об'єкта S_i ; α_{ij} — мітка, що характеризує належність об'єкта S_i класу K_j ($j = \overline{1, l}$). Тут $\alpha_{ij} \in \{0, 1, \Delta\}$, де 1 означає, що об'єкт S_i належить класу об'єкта K_j , а 0 — протилежне, Δ позначено ситуацію, коли алгоритм A не може визначити належність об'єкта S_i жодному із заданих класів.

Основне завдання полягає в обчисленні за інформацією I_0 та описом $I(S)$ допустимого об'єкта S значення елементарних предикатів $P_j(S_i)$, $S_i \in K_j$.

Інакше кажучи, потрібно побудувати такий алгоритм A , що

$$A(I_0, I(S)) = (\alpha_1^A(S), \alpha_2^A(S), \dots, \alpha_l^A(S)), \quad (1)$$

де $\alpha_j^{A(S)} \in \{0, 1, \Delta\}$, $j=1, \dots, l$.

На множині $\{A\}$ таких алгоритмів задано функціонал якості $\varphi(A)$ алгоритму A . Уточнена задача полягає в наступному: серед алгоритмів, визначених в (1), знайти алгоритм A^* такий, що

$$\varphi(A^*) = \sup_{A \in \{A\}} \varphi(A). \quad (2)$$

Інакше кажучи, серед алгоритмів розпізнавання потрібно вибрати алгоритм, що найбільш точно розпізнає об'єкти із контрольної вибірки.

Задача побудови інтегратора з урахуванням вибраного правила інтеграції полягає у наступному.

Розглянемо деяке правило інтеграції E , що залежить від вектора параметрів Θ і що дає змогу перетворити виходи m локальних класифікаторів (списки кандидатів Y_1, \dots, Y_m) у розв'язок Z . Тут $Y_j = (y_1^{(j)}, \dots, y_l^{(j)})$, де $y_i^{(j)}$ — оцінка близькості об'єкта S_j до класу K_i , отримана j -м класифікатором ($j=1, m$; $i=1, l$): $Z = E(Y_1, \dots, Y_m, \Theta)$.

Задача побудови інтегратора полягає у виборі прийнятної моделі E та оптимізації її параметрів Θ , що забезпечують найкращу якість розпізнавання

$$\Theta^* = \arg \max \{\varphi(\Theta^*)\}. \quad (3)$$

РОЗВ'ЯЗАННЯ ЗАДАЧІ

Для розв'язання задачі (1) пропонується модель локальних класифікаторів для розпізнавання зображень компонентів особи, яка ґрунтується на оцінці взаємозв'язку між ознаками компонента особи, що розпізнається. Далі наведено етапи розв'язання задачі запропонованою моделлю локальних класифікаторів.

Виокремлення сильнозв'язаних наборів ознак зображень компонентів обличчя. На цьому етапі визначаються n^* «незалежні» підмножини ознак компонентів обличчя.

Розглянемо всі можливі підмножини, що не перетинаються, з набору символів $\{x_1, \dots, x_p, \dots, x_n\}$. Позначимо W загальну множину таких підмножин. З множини W виокремимо n^* «незалежних» підмножин, що складаються з сильнозв'язаних ознак. Ці підмножини є загальною множиною $W_{\text{нр}} (W_{\text{нр}} \in W)$, що визначається $W_{\text{нр}} = \{\Omega_1, \dots, \Omega_q, \dots, \Omega_{n^*}\}$. При цьому елементи $\Omega_q (q=1, n^*)$ залежать як від n^* , так і від початкової інформації. Тому вони будуть різними для кожного класу $K_j (j=1, l)$.

Формування набору ідентифікаційних ознак зображень компонентів особи. На цьому етапі формується набір n^* ідентифікаційних ознак зображення компонента особи.

Для формування набору ідентифікаційних ознак зображення кожного компонента підмножини із Ω_q виокремимо одного типового представника $\psi_{i_q} (\psi_{i_q} \in \Omega_q)$. Сукупність таких представників визначає набір ідентифікаційних ознак, що формують простір розмірності n^* . При цьому n^* набагато менше розмірності вихідного простору ознак $n (n^* \ll n)$. Далі простір ідентифікаційних ознак позначимо Ψ^* : $\Psi^* = (\psi_{i_1}, \dots, \psi_{i_q}, \dots, \psi_{i_{n^*}})$.

Визначення функції відмінності $d(F_u, F)$ між зображеннями компонентів обличчя F_u та F . На цьому етапі визначається функція відмінності зображень компонентів обличчя F_u та F , що описує міру відмінності між зображеннями і задається у вигляді

$$d(F_u, F) = \sum_{q=1}^{n^*} \lambda_q (a_{ui_q} - a_{i_q})^2, \quad (4)$$

де a_{ui_q} — значення i_q -ідентифікаційної ознаки зображення F_u , λ_q — невідомий параметр, що характеризує важливість відповідної ознаки зображення компонента обличчя.

Визначення узагальненої функції відмінності між зображенням компонента обличчя F і класом K_j . На цьому етапі визначається узагальнена функція відмінності, що описує відмінність зображення компонента обличчя F від класу K_j . Узагальнена функція відмінності має вигляд

$$D(K_j, F) = \sum_{F_u \in K_j^*} \gamma_u d(F_u, F), \quad (5)$$

де γ_u — параметр, що характеризує важливість зображення F_u у навчальній вибірці.

Визначення функції близькості $R(K_j, F)$ між зображенням компонента обличчя F і класом K_j . Тут функція близькості між зображенням компонента обличчя F та класом K_j визначається за допомогою радіальних функцій

$$R(K_j, F) = \frac{1}{1 + \tau D(K_j, F)}, \quad (6)$$

де τ — параметр алгоритму розпізнавання. Для зображення компонента обличчя F з використанням радіальних функцій на основі виразу (6) можна обчислити оцінки належності $F: R(K_1, F), \dots, R(K_j, F), \dots, R(K_l, F)$. Ця послідовність оцінок власності визначає «м'які» розв'язки локального класифікатора.

Таким чином, наведені етапи формують модель локальних класифікаторів, основу на оцінці взаємозв'язку між ознаками розглянутого компонента обличчя, що розпізнається.

Розв'язання оптимізаційної задачі (2) щодо параметрів n^* , Ω_q , γ_u , τ дає змогу для зазначеної моделі визначити конкретний локальний класифікатор, який відповідає компоненту обличчя, що розпізнається. Розглянемо розв'язання задачі побудови інтегратора локальних розв'язків, сформульованої у вигляді (3). Існує низка технологій (правил) комбінування класифікаторів на рівні прийняття рішень, які згідно з [3] ґрунтуються на трьох підходах: абстрактному, ранговому та бальному.

У випадку першого підходу на вхід класифікатора надходить по одній мітці класу від кожного класифікатора, у разі другого підходу кожен класифікатор дає кілька міток, ранжованих від найімовірніших до менш імовірних, у випадку третього підходу кожен класифікатор виводить n^* кращих міток з оцінками достовірності. Зазначені підходи до комбінування класифікаторів поєднує загальна ідея — інтегрувати інформацію після того, як кожен із класифікаторів представить свої локальні розв'язки. Обґрунтованість такої інтеграції полягає в тому, що вона дає змогу об'єднати сукупність класифікаторів, деякі з яких можуть бути слабкими. Мета інтеграції класифікаторів — забезпечити необхідний рівень продуктивності та якості розпізнавання.

Аналіз публікацій з цієї галузі досліджень, наприклад [5–9], свідчить про недостатню розвиненість теорії побудови методів інтеграції класифікаторів та

аналізу їхньої ефективності. У [5] наведено основні задачі (для їхнього розв'язання потрібно проведення цілеспрямованих досліджень), а саме:

— задачі з двома класами і низькорозмірні ознакові простори (багато методів інтеграції класифікаторів дієві за певних обмежень, що накладаються, зокрема, на кількість класів, а також на кількість ознак);

— теоретичне доведення, що інтеграція слабких класифікаторів або інтеграція слабких і сильних класифікаторів може спричинити підвищення точності та продуктивності інтегратора;

— інтеграція великої кількості класифікаторів (один із можливих підходів полягає в організації ієрархічної структури класифікаторів, у якій розв'язок формується поступово, починаючи з класифікаторів нижнього рівня та закінчуючи класифікаторами верхнього рівня);

— об'єднання класифікаторів за наявності великої кількості класів (наприклад, понад 500).

Розглянута в цій роботі багаторівнева система розпізнавання осіб має низку можливостей, які надають змогу зменшити вплив зазначених під час вибору правила комбінування локальних класифікаторів. По-перше, це зниження розмірності ознакового простору за рахунок використання локальних класифікаторів, кожен з яких розпізнає конкретний компонент обличчя, що дає змогу опустити операцію конкатенації ознак компонентів обличчя, що розглядається, використання якої спричиняє формування єдиного вектора ознак для всієї області обличчя, що має велику розмірність. По-друге, це використання обмеженої кількості класифікаторів, що визначається невеликою кількістю компонентів, які використовують для розпізнавання обличчя.

Зазначимо, що вибране правило інтеграції класифікаторів має забезпечити розв'язання задачі розпізнавання особи за наявності чималої кількості класів (наприклад, понад 500), що виникає в процесі пошуку великої бази облич. Як зазначено в [5], нейронні мережі можуть успішно адаптуватися до такої ситуації, що зумовлює доцільність їхнього використання як інтегратора в розглянутій дворівневій системі розпізнавання обличчя. Під час використання простої нейронної мережі [12] в цій системі вихідні значення m локальних класифікаторів передаються $m \times l$ нейронам вхідного шару нейронної мережі, де l — кількість заданих класів. У вихідному шарі інтегратора кожен з l нейронів генерує оцінку належності зображення обличчя, що розпізнається, одному із заданих класів. Згодом на основі цих оцінок з використанням функції Softmax формується остаточне рішення щодо обличчя, яке розпізнається. Для навчання нейронної мережі було вибрано алгоритм зворотного поширення. Як функцію помилки в процесі навчання нейромережевої моделі використовується категоріальна кросентропійна втрата на навчальних вибірках Loss [9]:

$$\text{Loss} = -\sum_{j=1}^l z_j \times \log(z_j^*), \quad (7)$$

де z_j^* — вихідне значення, відповідне j -класу на виході моделі (значення прогнозу), z_j — раніше відоме значення, відповідне j -класу, l — кількість класів, $j = \overline{1, l}$.

Остаточне рішення формується на основі оцінок імовірностей, отриманих на виході нейромережевого інтегратора з використанням функції Softmax, яка є функцією виходу, рекомендованою для використання з функцією категоріальної кросентропійної помилки.

ЕКСПЕРИМЕНТ З РОЗПІЗНАВАННЯ ОБЛИЧ ОСІБ

Для проведення експериментів використовувалася база фронтальних зображень осіб FERET [10], що включає 1187 зображень облич 69 осіб. Розмір кожного зображення становив за шириною 256 пікселів і за висотою 384 пікселя.

Таблиця 1. Якість різних варіантів інтеграції локальних класифікаторів

| Інтегратор | Точність розпізнавання | | |
|----------------|---------------------------------|---------------------------------|-------------------------------------|
| | Перша область обличчя (очі+ніс) | Друга область обличчя (ніс+рот) | Третя область обличчя (очі+ніс+рот) |
| Нейромережевий | 0.951 | 0.895 | 0.973 |
| Баєсовський | 0.923 | 0.886 | 0.945 |
| Лог-лінійний | 0.938 | 0.891 | 0.954 |

Для дослідження ефективності різних моделей класифікаторів було проведено експерименти з розпізнавання осіб з урахуванням покомпонентного підходу. Тому для проведення експериментів на кожному зображенні особи із 690 зображень осіб, вибраних із бази FERET (по 10 зображень осіб із 69 осіб), з використанням алгоритму, що реалізує модифікований метод Віюлі-Джонса [15], було виокремлено локальні області, кожна з яких представляла розглядуваний компонент обличчя: праве око, ліве око, ніс та рот. Зображення кожного компонента були включені у відповідну вихідну вибірку. Таким чином, було сформовано чотири типи вихідних вибірок, кожна з яких включала 690 зображень відповідного компонента особи. Як ознаки розпізнаваних компонентів особи використовувалися дескриптори, що визначаються за допомогою методу локальних бінарних шаблонів, докладно описаних у [15].

Експерименти проводилися з використанням розглянутого нейромережевого інтегратора, баєсовського інтегратора [6] та лог-лінійного інтегратора [3]. Результати реалізації цих інтеграторів наведено у табл. 1.

Як видно з табл. 1, нейромережевий інтегратор має кращу якість, ніж інші два інтегратори. При цьому якість лог-лінійного інтегратора вища, ніж у баєсовського, завдяки чому останній орієнтований на інтеграцію незалежних класифікаторів. Враховуючи, що локальні класифікатори, які інтегруються, є залежними, оскільки сформовані в рамках однієї моделі, ця обставина і вплинула на якість баєсовського інтегратора. Більш висока якість розпізнавання першої області особи, що включає очі і ніс, порівняно з другою областю, що включає ніс і рот, підтверджує факт, що верхня частина обличчя є більш інформативною ніж нижня.

ВИСНОВКИ

Наявність проблеми створення надійних систем розпізнавання стимулювала дослідників пропонувати нові підходи до комбінування мультикласифікаторів. Для реалізації цих підходів необхідно досліджувати адаптивність комбінації, визначати рівні комбінації та типи порогових значень під час прийняття проміжних та остаточних рішень. У статті показано, що методи комбінування класифікаторів можна успішно застосовувати для задач розпізнавання осіб.

У роботі основну увагу приділено постановці та розв'язанню задач побудови моделі локальних класифікаторів та вибору правила їхнього комбінування. Запропонована модель локальних класифікаторів ґрунтується на переході від вихідної системи залежних ознак великої розмірності до системи незалежних ідентифікаційних ознак суттєво меншої розмірності. Достатньо висока якість розпізнавання поєднань основних компонентів обличчя (очі, ніс, рот) свідчить про збереження дієвості розглянутої у статті багаторівневої системи розпізнавання осіб за наявності часткової оклюзії. Ця особливість зазначеної системи розширює сферу її практичного застосування, зокрема, в системах активного відеоспостереження бортових систем [13].

СПИСОК ЛІТЕРАТУРИ

1. Кийко В.М., Кийко К.В., Недашківський О.В., Мацелло В.В., Шлезінгер М.І. Метод компьютерной идентификации человека по изображению лица: пат. № 81514 Україна, заявл. 30.01 2006; опубл. 10.01.2008, Бюл. № 1. 26 с.
2. System and method for identifying object in video: United States Patent: № US 8,064,641 B2; Inventors: Anchishkin Y., Kyuko V., Musatenko Y., Schlesinger M., Kovtun I., Milshteyn K. Date of Patent: Nov. 22, 2011.
3. Method and system of person identification by facial image: United States Patent: № US 8,184,914 B2; Inventors: Tsariov A., Musatenko Y., Schlesinger M., Kyuko V., Kyiko K., Nedashkyvskiy O.; Date of Patent: May 22, 2012.
4. Gumede A., Viriri S., Gwetu M. Hybrid component-based face recognition. *Proc. Conference on Information Communication Technology and Society (ICTAS)*. (Durban, South Africa, 2017). P. 1–6. <https://doi.org/10.1109/ICTAS.2017.7920665>.
5. Adam H., Serestina V. Invariant Feature Extraction for Component-based *Facial Recognition*. *International Journal of Advanced Computer Science and Applications*. 2020. Vol. 11. <https://doi.org/10.14569/IJACSA.2020.0110386>.
6. Paul S.K., Bouakaz S., Rahman C.M. et al. Component-based face recognition using statistical pattern matching analysis. *Pattern Anal Applic* 24, 2021. P. 299–319. <https://doi.org/10.1007/s10044-020-00895-4>.
7. Камиллов М.М., Фазылов Ш.Х., Мирзаев Н.М., Раджабов С.С. Модели алгоритмов распознавания, основанных на оценке взаимосвязанности признаков. Ташкент: Наука и технологии, 2020. 148 с.
8. Mohandes M., Deriche M., Aliyu S.O. Classifiers combination techniques: *A comprehensive review*. *IEEE Access*. 2018. Vol. 6. P. 19626–19639. <https://doi.org/10.1109/ACCESS.2018.2813079>.
9. Opanasenko V.N., Kryvyi S.L. Synthesis of neural-like networks on the basis of conversion of cyclic hamming codes. *Cybernetics and Systems Analysis*. 2017. Vol. 53, N 4. P. 627–635. <https://doi.org/10.1007/s10559-017-9965-z>.
10. Tulyakov S., Jaeger S., Govindaraju V., Doermann D. Review of classifier combination methods. *Machine Learning in Document Analysis and Recognition*. Berlin: Springer, 2008. P. 361–386.
11. De Stefano C., Fontanella F., di Freca A.S. A novel naive Bayes voting strategy for combining classifiers. *Proc. Int. Conf. Front. Hand-writing Recognit.* (Bari, Italy, Sep. 2012). P. 467–472.
12. Lan Y.-D., Gao L. A new model of combining multiple classifiers based on neural network. *Proc. 4th Int. Conf. Emerg. Intell. Data Web Technol.* 2013. Vol. 2. P. 154–159.
13. Chitroub S. Classifier combination and score level fusion: Concepts and practical aspects. *Int. J. Image Data Fusion*. 2010. Vol. 1, N 2. P. 113–135.
14. Color FERET Database. NIST. URL: <https://www.nist.gov/itl/products-and-services/color-feret-database>.
15. Viola P., Jones M.J. Robust real-time face detection. *International Journal of Computer Vision*. 2004. Vol. 57. P. 137–154.
16. Ojala T., Pietikainen M., Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*. 2002. Vol. 24, N 7. P. 971–987.
17. Opanasenko V., Palahin A., Zavyalov S. The FPGA-based problem-oriented on-board processor. *Proc. of the 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, (IDAACS'2019)*. (Metz, France, 18–21 Sept. 2019). 2019. Vol. 1. P. 152–157. <https://doi.org/10.1109/IDAACS.2019.8924360>.

V.M. Opanasenko, Sh.Kh. Fazilov, S.S. Radjabov, Sh.S. Kakharov

MULTILEVEL FACE RECOGNITION SYSTEM

Abstract. The problem of biometric person identification based on component-based face recognition is considered. It is shown that the face recognition system can be represented as a hierarchically organized multilevel system in which an ensemble of local classifiers forms “soft” decisions about the belonging of images of individual components of the face to given classes. Then, based on the integration of these decisions, the formation of the final decision on whether the recognized face belongs to one of the given classes is realized. The problems of constructing a model of a local classifier, as well as choosing an integrator of intermediate solutions of local classifiers, are formulated and solved.

Keywords: pattern recognition, multilevel recognition system, classifier ensemble, classifier combination rule, decision making.

Надійшла до редакції 13.02.2023