

УДК 004.8:519

DOI: 10.31673/2412-9070.2020.02933

І. С. ЩЕРБИНА, канд. техн. наук, доцент;  
Д. В. ЛАСКАВИЙ, магістрант,  
Державний університет телекомунікацій, Київ

## ПОРІВНЯЛЬНА ХАРАКТЕРИСТИКА СУЧАСНИХ МЕТОДІВ РОЗПІЗНАВАННЯ ЛЮДЕЙ НА ОСНОВІ КОМП'ЮТЕРНОГО ОБРОБЛЕННЯ ЇХ ЗОБРАЖЕННЯ

*Проблема автоматизованого розпізнавання облич та жестів є відносно новою і досі повністю не вирішеною. Останніми роками було запропоновано низку різних методів і алгоритмів для оброблення, локалізації та розпізнавання облич і жестів у статичних зображеннях, таких як «власні обличчя» (аналіз основних компонентів, PCA), нейронні мережі, еволюційні алгоритми, алгоритм AdaBoost, метод підтримання вектора тощо. Однак ці підходи до розпізнавання об'єктів мають недостатню точність, надійність та швидкість у складних реальних умовах, що характеризуються наявністю шуму в зображеннях та відеопослідовностях.*

*У наш час системи відеоспостереження набули великого поширення. Є багато компаній, які виробляють своє обладнання для цієї мети. Це картки для захоплення відео зі спеціальним програмним забезпеченням. Наприклад, для того, щоб система спостереження в пункті пропуску підрахувала кількість людей, що проходять, необхідно витратити значну суму грошей для її програмного забезпечення, оскільки буде потрібне спеціальне програмне оброблення відеопотоку, яка розпізнаватиме зображення та обчислюватиме їх.*

*Слід зазначити, що існує низка факторів, які ускладнюють розпізнавання об'єктів у статичних зображеннях та відеопослідовностях. До них належать: зміна умов освітлення під час процесу зйомки, низька якість відеозображень, труднощі відокремлення об'єкта від фону, наявність багатьох об'єктів у відеокадри тощо.*

*У статті визначено завдання виявлення об'єктів на зображенні, а також методи оброблення та аналізу даних. Вивчено методи розпізнавання, які є одними з перших практичних завдань, що стало стимулом для розвитку теорії розпізнавання об'єктів. Проаналізовано питання розпізнавання облич та жестів, що знаходить застосування у різних сферах людської діяльності. Розглянуто алгоритми розпізнавання облич.*

**Ключові слова:** розпізнавання облич; штучна нейронна мережа; штучний інтелект.

### ВСТУП

Завдання виокремлення особи людини природною для подальшої ідентифікації завжди перебувало серед найбільш пріоритетних завдань для дослідників, що працюють у сфері систем машинного зору і штучного інтелекту. Проте багато досліджень, що виконуються в провідних наукових центрах усього світу протягом останніх кількох десятиліть, так і не привело до створення реально працюючих систем комп'ютерного зору, здатних виявляти і розпізнавати людину за будь-яких умов [1; 2]. Незважаючи на близькість завдань і методів, використовуваних під час розроблення альтернативних систем біометричної ідентифікації людини, зокрема ідентифікації за відбитками пальців або за малюнком райдужної оболонки, системи ідентифікації за зображенням особи істотно поступаються зазначеним системам.

Серйозною проблемою, що стоїть перед системами комп'ютерного зору, є велика мінливість візуальних образів, пов'язана зі змінами освітленості, забарвлення, масштабів, ракурсів спостереження. Окрім того, люди перебувають на вулицях і в приміщенні одягненими, що призводить до істотної мінливості у зображеннях однієї і тієї самої людини. Однак найбільш складним завданням комп'ютерного зору є проблема усунення неоднозначності, що виникає під час проєкціювання тривимірних об'єктів реального світу на плоскі

зображення. Колір і яскравість окремих пікселів на зображенні також залежать від великої кількості важко прогнозованих факторів, до яких належать:

- кількість і розташування джерел світла;
- колір та інтенсивність випромінювання;
- тіні або віддзеркалення від навколишніх об'єктів.

Завдання виявлення об'єктів на зображенні ускладнюється в разі великого обсягу даних, що містяться в зображенні. Зображення може мати тисячі пікселів, кожний із яких дуже важливий. Повне використання інформації, що міститься в зображенні, потребує аналізу кожного пікселя на приналежність його об'єкту або фону з урахуванням можливої мінливості об'єктів. Такий аналіз може призвести до чималих витрат у необхідній пам'яті і продуктивності комп'ютера [3; 4].

Розв'язання цієї проблеми полягає в правильному виборі опису об'єктів, для виявлення і розпізнавання яких створюється система. Опис об'єкта має враховувати найбільш характерні його особливості і бути добре поданим, щоб відрізнити даний об'єкт від інших елементів навколишньої сцени. Щоб уникнути суб'єктивності під час вибору потрібного опису, можна скористатися методами автоматичного вибору відповідних характеристик об'єкта, які реалізуються в генетичних алгоритмах і в процесі навчання штучних нейронних

© І. С. Щербина, Д. В. Ласкавий, 2020

мереж. Водночас існує низка параметрів в описі об'єкта, які в даний час має вибрати дослідник, що розробляє систему виявлення і розпізнавання. До такого вибору належать:

- вибір між 2D і 3D-виставою сцени і об'єкта. Алгоритми, що використовують 2D-уявлення, зазвичай простіші, ніж 3D-алгоритми, але потребують чимало різних описів, відповідних поданням об'єкта за різних умов спостереження;
- вибір між описом об'єкта як єдиного цілого або як системи, що складається з певної кількості взаємозв'язаних елементів;
- вибір між системою ознак, що ґрунтуються на геометричних чи інших ознаках та описують характеристику специфіки об'єкта.

У найзагальнішому випадку алгоритм вирішення задачі виявлення та ідентифікації людини за зображенням його особи складається з таких очевидних кроків:

- виявлення факту присутності людини на аналізованій сцені;
- виокремлення фігури людини;
- виокремлення голови;
- визначення ракурсу спостереження голови (анфас, профіль);
- виокремлення особи;
- порівняння з еталонами та ідентифікація.

**Аналіз досліджень і публікацій.** У роботі [5] розглянуто розпізнавання осіб на групових фотографіях із упровадженням алгоритмів сегментації. Рекомендації, надані в цій роботі, мають здебільшого загальний характер. Їх використання потребує нових нетрадиційних підходів до розв'язання проблеми в цілому і є підставою на застосування методів штучного інтелекту.

Сьогодні, як зазначається в [6], відеонагляд є одним із найпопулярніших технологічних інструментів упровадження безпеки. При цьому недостатньо мати лише систему відеонагляду, а й алгоритми, які в автоматизованому вигляді зможуть аналізувати все, що потрапляє в поле зору, та виконувати будь-які дії стосовно отриманих даних.

Як пояснюється в літературі іноземного автора Taigman Y, серед усіх можливих об'єктів для систем розпізнавання найбільшу зацікавленість становлять зображення облич.

У фундаментальних працях [8; 9] досліджено розпізнавання жестів на відеопослідовності в режимі реального часу на основі застосування методу Віюлі-Джонса. Проте якість розроблених алгоритмів розпізнавання як динамічних, так і статичних, із використанням кольорових відеокамер і тривимірних сенсорів все ще залишається недостатнім для побудови практичних систем людино-машинної взаємодії.

**Мета дослідження.** Метою дослідження є аналіз існуючих алгоритмів розпізнавання, порівняльна

характеристика та отримання результатів дослідження на основі наявної інформації.

Для досягнення поставленої мети необхідно розв'язати такі задачі:

- відшукування сучасних методів і способів розпізнавання;
- обґрунтування вибраних вирішень на основі аналізу отриманих результатів.

## ОСНОВНА ЧАСТИНА

### Методи, засновані на геометричних характеристиках обличчя

Метод розпізнавання за геометрією особи [10] складається з трьох компонент:

1) перетворення вихідного зображення в початкове уявлення (може включати в себе як передоброблення, так і математичні перетворення, наприклад обчислення головних компонент);

2) виокремлення ключових характеристик (наприклад, беруться перші  $n$  головних компонент або коефіцієнтів дискретного косинусного перетворення);

3) створення механізму класифікації (моделювання): кластерна модель, метрика, нейронна мережа тощо.

Окрім цього, побудова методу розпізнавання ґрунтується на апріорній інформації про предметну область (у даному випадку — характеристики особи людини) і коригується експериментальною інформацією, що з'являється у процесі розроблення методу. Завдання ідентифікації особистості з використанням геометрії особи конкретизується як необхідність реалізації модуля, що дає змогу формувати і заносити дані про користувачів у базу даних (модуль реєстрації користувача), і модуля, що виконує ідентифікацію (модуль ідентифікації).

До розроблюваних модулів висуваються такі вимоги:

- крос-платформність;
- можливість взаємозалежного використання модулів;
- висока відмовостійкість;
- простота в роботі.

Для досягнення поставленої мети необхідно розв'язати такі завдання:

- проаналізувати наявні методи ідентифікації з використанням оброблення геометрії особи;
- вибрати оптимальний метод для створення даного модуля;
- розробити алгоритм ідентифікації особистості за зображенням, яке містить її обличчя;
- розробити базу даних для зберігання інформації про користувачів і використання ними зазначених модулів;
- розробити програмний модуль, який буде вносити дані в еталонну вибірку.

Модуль реєстрації складається з таких частин:

- база даних для зберігання даних про користувача;
- модуль редагування даних для зміни даних про користувачів;
- модуль оброблення зображення для пошуку особи користувача на пред'явлені зображення.

Модуль реєстрації має виконувати такі функції:

- оброблення даних;
- створення нового запису користувача;
- зміну запису користувача;
- видалення запису;
- оброблення зображення;
- попереднє оброблення вихідного зображення;
- пошук особи на зображенні;
- збереження зображень.

Схему алгоритму функціонування модуля реєстрації користувача наведено на рис. 1.

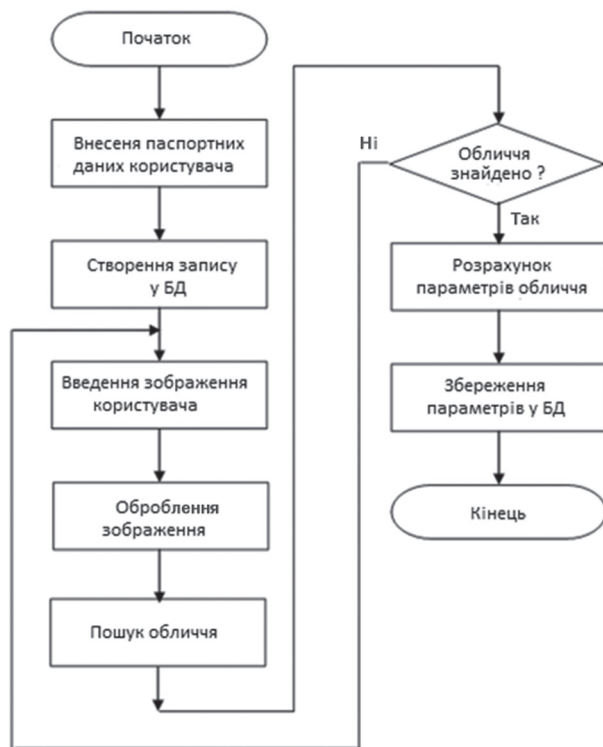


Рис. 1. Схему алгоритму функціонування модуля реєстрації користувача

### Порівняння шаблонів

Порівняння шаблонів (*Template Matching*) полягає у виокремленні ділянок особи на зображенні (рис. 2), з подальшим порівнянням цих ділянок для двох різних зображень. Кожна збіжна ділянка збільшує ступінь схожості зображень. Це також один із історично перших методів розпізнавання людини за зображенням особи. Для порівняння ділянок застосовуються найпростіші алгоритми, на кшталт попиксельного порівняння [11; 12].

Недолік цього методу полягає в тому, що він потребує багато ресурсів як для зберігання ділянок, так і для їх порівняння. З огляду на те, що вико-

ристовується найпростіший алгоритм порівняння, зображення мають бути зняті в строго встановлених умовах: неприпустимо помітні зміни ракурсу, освітлення, емоційного вираження тощо.



Рис. 2. Ділянки, що входять у шаблон особи

### Багатошарові нейронні мережі

Архітектура багатошарових нейронних мереж (БНМ) характеризується послідовно з'єднаними шарами, де нейрон кожного шару своїми входами пов'язаний із усіма нейронами попереднього шару, а виходами — наступного. Двошарова НМ може з будь-якою точністю виконати апроксимацію на будь-яку багатовимірну функцію. Формування лінійних розподільвальних поверхонь є важливим завданням цієї мережі та унеможливає використання такої логічної функції, як «виняткове ЧИ». Такі НМ здатні формувати опуклі поверхні у просторі вирішень у разі використання двох вирішальних шарів. А якщо застосовується три вирішальних шари, то формуються ділянки будь-якої складності — опуклі чи неопуклі. БНМ при цьому не втрачає своєї підсумовувальної здатності. БНМ навчаються за допомогою методу градієнтного спуску в просторі ваг з метою, аби мінімізувати кількість помилок мережі, тобто оберненого методу поширення помилки. При цьому коригування ваги поширюється в зворотному напрямі через ваги, що з'єднують нейрони від входу до виходу. Найпростіший приклад — це коли мережа може відновлювати вхідні зображення. Вона реконструює це зображення і вираховує якість, тим самим оцінюючи, наскільки мережа розпізнала це вхідне зображення. Перевагою цього методу є те, що мережа здатна відновити навіть зашумлену картинку, але вона не придатна для більш серйозних завдань. БНМ може класифікувати будь-яке зображення, оскільки вхідними даними є набір характеристик зображення, а на виході маємо активний нейрон, який зазначить належність до класу, що має бути розпізнано (рис. 3). Активність нейрона має бути вищою за певний поріг. Якщо цей результат інакший, вважатимемо що шуканий образ не належить до жодного класу, які відомо. Навчання характеризується процесом

добору на вхід класів, та встановлює належність до певного з них. Такий процес називається «навчанням із учителем». Його зручно застосовувати для розпізнавання людини за зображенням особи або групи осіб, оскільки забезпечується безпосереднє порівняння з тими прикладами, що є в базі. Для таких завдань потрібно використовувати набір ключових характеристик, завдяки яким здійснюється пошук [4].

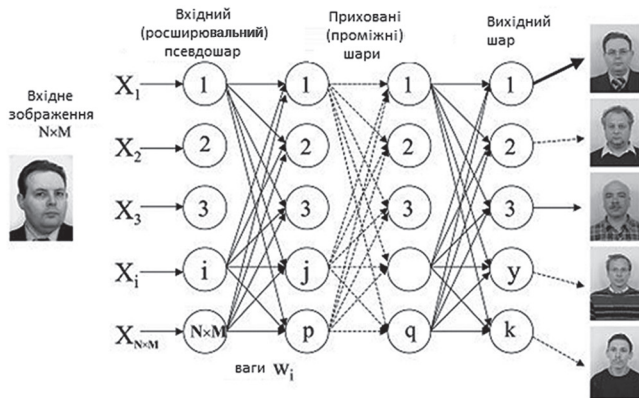


Рис. 3. Багатошарова нейронна мережа для класифікації зображень

За допомогою БНМ можна виконати розподілення зображень за такими параметрами, як відстань між вибраними частинами обличчя, наприклад ніс, рот або очі. У класичному вигляді БНМ, її нейронні з'єднання подаються як одновимірний вектор, не зважаючи на те, що він є двовимірним. БНМ має стійкість до змін масштабу, зсуву, поворотам або спотворень.

Ще є марковська модель, яка об'єднується з багатошаровою моделлю, тобто є гібридною.

### Результати дослідження

У тестах застосовувались алгоритми, засновані на гнучкому порівнянні на графах і деякі модифікації методу головних компонент PCA. Кожний алгоритм мав приблизно однакову ефективність. З огляду на це важко або навіть неможливо визначити чіткі відмінності між ними (особливо якщо узгодити дати тестування). Точність розпізнавання фронтальних зображень, зроблених в один день, становила, як правило, 95%. Тоді як точність фото, які було зроблено різними пристроями за різного рівня освітлення, зазвичай становить 80%. Для зображень, виконаних із різницею в рік, точність розпізнавання досягала майже 50%. Зауважимо, що навіть 50% — це більш ніж прийнятна точність роботи системи подібного роду.

### Обговорення результатів дослідження

Багатошарові мережі мають досить потужні можливості. Наприклад, двошарову мережу, в якій перший шар містить сигмоїдальну, а другий шар — лінійну функцію активації, можна навчи-

ти апроксимувати з довільною точністю будь-яку функцію зі скінченною кількістю точок розриву. Тут можна сформулювати наступні висновки. Вхід функції активації нейрона визначається зміщенням і сумою зважених входів. Вихід нейрона залежить як від входів нейрона, так і від виду функції активації. Один нейрон не може вирішувати складні завдання, проте кілька нейронів, об'єднаних в один або кілька шарів, мають більші можливості.

Архітектура мережі складається з опису того, скільки шарів має мережа, кількості нейронів у кожному шарі, виду функції активації кожного шару та інформації про з'єднання шарів. Також архітектура мережі залежить від тієї конкретної задачі, яку вона має розв'язувати.

Робота мережі полягає в обчисленні виходів мережі на основі відомих входів із метою формування бажаного відображення вхід/вихід. Конкретне завдання визначає кількість входів і кількість виходів мережі. Окрім кількості нейронів у вихідному шарі мережі для проектувальника важлива кількість нейронів у кожному шарі. Більша кількість нейронів у прихованих шарах забезпечує більш потужну мережу. Якщо має бути реалізовано лінійне відображення, то слід використовувати нейрони з лінійними функціями активації.

При цьому треба пам'ятати, що лінійні нейронні мережі не можуть формувати нелінійні відображення. Використання нелінійних функцій активації дає можливість налаштувати нейронну мережу на реалізацію нелінійних зв'язків між входом і виходом.

### ВИСНОВКИ

1. Розглянуто алгоритми для розпізнавання об'єктів на статичних зображеннях і динамічних об'єктах щодо відеопослідовності в режимі реального часу.

2. Викладено результати застосування розроблених алгоритмів стосовно розв'язання задачі розпізнавання об'єктів на статичних зображеннях і рухомих об'єктах на відеопослідовності в режимі реального часу.

3. Для більшості сучасних систем автоматичного розпізнавання осіб основним завданням є порівняння даного зображення особи з набором зображень осіб із бази даних. Характеристики систем автоматичного розпізнавання осіб у цьому разі оцінюються через визначення ймовірностей помилкової відмови в розпізнаванні (для зображення особи, наявної в базі, ухвалюється рішення як про невпізнану особу) і помилкового розпізнавання.

4. Застосовано оцінку стійкості до збурення зображень, спричинене комбінацією зі складними фонами, мінливістю освітлення, зміною зачіски та іншими факторами.

**Список використаної літератури**

1. **Deepface: Closing the gap to human-level performance in face verification** / Y. Taigman [et al.] // *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2014. С. 1701–1708.
2. **Schroff F., Kalenichenko D., Philbin J. Facenet: A unified embedding for face recognition and clustering** // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015. С. 815–823.
3. **Лукьяниці А. А., Шишкин А. Г.** Цифрова обробка відеоізображень // *Ай-Ес-Ес Прес*, 2009. С. 511.
4. **Журавльов Ю. І.** Розпізнавання: математическі методи // *Фазис*. 2006. С. 147.
5. **Deep Face Recognition** / O. M. Parkhi [et al.] // *BMVC*. 2015. Т. 1. № 3. С. 6.
6. **A light CNN for deep face representation with noisy labels** / X. Wu [et al.] // *arXiv preprint arXiv:1511.02683*. 2015.

7. **Amos B., Ludwiczuk B., Satyanarayanan M.** *Openface: A general-purpose face recognition library with mobile applications* // *CMU School of Computer Science*. 2016.

8. **Болотова Ю. А.** Алгоритмы обработки и анализа изображений иерархической временной сетью: дис. на соискание уч. степени канд. техн. наук. Томск, 2013.

9. **Фан Н. Х., Буй Т. Т. Ч., Спицын В. Г.** Распознавание жестов на видеопоследовательности в режиме реального времени на основе применения метода Виолы-Джонса, алгоритма CAMShift, вейвлет-преобразования и метода главных компонент // *Вест. Томск. гос. ун-та*. 2013. Т. 23. № 2. С. 102–111.

10. **Распознавание** лиц на групповых фотографиях с использованием алгоритмов сегментации / А. И. Шерстобитов, В. П. Федосов, В. А. Приходченко, Д. В. Тимофеев // *Известия ЮФУ: технические науки*. 2013. №5.

И. С. Щербина, Д. В. Ласкавый

**СРАВНИТЕЛЬНАЯ ХАРАКТЕРИСТИКА СОВРЕМЕННЫХ МЕТОДОВ РАСПОЗНАВАНИЯ ЛЮДЕЙ  
НА ОСНОВЕ КОМПЬЮТЕРНОЙ ОБРАБОТКИ ИХ ИЗОБРАЖЕНИЯ**

Проблема автоматического распознавания лиц и жестов является относительно новой и еще не полностью решенной. В последние годы был предложен ряд различных методов и алгоритмов для обработки, локализации и распознавания лиц и жестов в статических изображениях, таких как «собственные лица» (анализ основных компонентов, PCA), нейронные сети, эволюционные алгоритмы, алгоритм AdaBoost, метод опорных векторов и т. д. Однако эти подходы для распознавания объектов имеют недостаточную точность, надежность и скорость в сложной реальной среде, характеризующейся наличием шума на изображениях и видеопоследовательностях.

В настоящее время системы видеонаблюдения стали очень широко распространенными. Есть очень много компаний, которые выпускают свое оборудование для этой цели. Это карты видеозахвата со специальным программным обеспечением. Например, чтобы система наблюдения на контрольно-пропускном пункте рассчитала количество проходящих людей, необходимо заплатить значительную сумму денег для ее реализации, поскольку потребуется специальная программная обработка видеопотока, которая программным путем будет распознавать изображения и рассчитывать их.

Следует отметить, что существует ряд факторов, которые усложняют распознавание объектов в статических изображениях и видеопоследовательностях. К ним относятся: изменение условий освещения в процессе съемки, низкое качество видеоизображений, сложность отделения объекта от фона, наличие множества объектов в видеокадре и т. д.

В статье определена задача обнаружения объектов на изображении, а также методы обработки и анализа данных. Изучены методы распознавания, которые являются одними из первых практических задач, что стало стимулом для развития теории распознавания объектов. Проанализированы вопросы распознавания лиц и жестов, которые находят применение в различных сферах человеческой деятельности. Рассмотрены алгоритмы распознавания лиц.

**Ключевые слова:** распознавание лиц; искусственная нейронная сеть; искусственный интеллект.

I. S. Shcherbina, D. V. Laskavyi

**COMPARATIVE CHARACTERISTICS OF MODERN METHODS OF RECOGNITION OF PEOPLE  
ON THE BASIS OF COMPUTER PROCESSING OF THEIR IMAGE**

The problem of automated recognition of faces and gestures is relatively new and has not yet been fully resolved. In recent years, a number of different methods and algorithms have been proposed for processing, localizing and recognizing faces and gestures in static images, such as «own faces» (Principal Component Analysis, PCA), neural networks, evolutionary algorithms, AdaBoost algorithm, support vector method, etc. However, these approaches for recognizing objects have insufficient accuracy, reliability, and speed in a complex real environment characterized by the presence of noise in images and video sequences.

Nowadays CCTV systems have become very widespread. There are so many companies that release their equipment for this purpose. These are video capture cards with special software. But there are problems, of course, this is software that is made for a certain range of tasks, in most cases it is very disturbing. For example, in order for the surveillance system at the checkpoint to calculate the number of passing people, it is necessary to pay a considerable amount of money, since it requires special software processing of the video stream in order to realize it in order to programmatically recognize the images and to calculate them.

It should be noted that there are a number of factors that complicate the recognition of objects in static images and video sequences. These include: changing lighting conditions during the shooting process, poor quality of video images, the difficulty of separating the object from the background, the presence of many objects in the video frame, etc. This article discusses the task of detecting objects in an image, as well as methods for processing and analyzing data. The recognition methods, which are one of the first practical tasks, are studied, which has become an incentive for the development of the theory of object recognition. The issues of recognition of faces and gestures, which finds application in various fields of human activity, are analyzed. In this paper, face recognition algorithms are considered.

**Keywords:** face recognition; artificial neural network; artificial intelligence.