

УДК 004.93

DOI <https://doi.org/10.32782/tnv-tech.2023.6.2>

АНАЛІЗ ТЕХНОЛОГІЙ ІДЕНТИФІКАЦІЇ ОБЛИЧЬ НА ЗОБРАЖЕННІ

Зінов'єва О. Г. – старший викладач кафедри комп'ютерних наук
Таврійського державного агротехнологічного університету
імені Дмитра Моторного
ORCID ID: 0000-0003-3760-8952

Стаття присвячена дослідженню та аналізу методів ідентифікації обличь на зображеннях. Розглядаються актуальні підходи та технології, що використовуються для розпізнавання осіб у великому обсязі даних. Зокрема, досліджуються методи машинного навчання, нейронні мережі та комп'ютерне бачення, які застосовуються для вдосконалення точності ідентифікації та розпізнавання осіб на зображеннях.

Ідентифікація обличь на зображенні – це завдання комп'ютерного зору, спрямоване на розпізнавання та класифікацію облич людей або інших об'єктів на фотографіях чи відео. Завдання розпізнавання осіб має безліч додатків в таких галузях, як біометрія, організація відеоконференцій, системи машинного зору в робототехніці, інтелектуальні системи безпеки і контролю доступу і т. п. Завдання виявлення обличчя на зображенні часто є першим кроком у процесі розв'язання задачі вищого рівня – розпізнавання обличчя, деталей обличчя або його міміки. Крім того, інформація про присутність та кількість осіб на зображенні може бути корисною в системах автоматичного обліку числа відвідувачів; системах пропускового контролю в установах, аеропортах та метро; автоматичних системах запобігання нещасним випадкам; інтелектуальних інтерфейсах «людина-комп'ютер»; у фототехніці для автоматичного фокусування на обличчі людини, а також для стабілізації зображення обличчя з метою полегшення розпізнавання емоцій; для розширення зони стереобачення під час створення систем 3D відображення. Аналіз методів ідентифікації облич на зображеннях відіграє важливу роль у розвитку технологій комп'ютерного зору та розпізнавання образів.

Стаття висвітлює основні виклики у цьому напрямку, такі як забезпечення стійкості до змін освітлення, позиції обличчя та зовнішніх факторів. Досліджуються також питання конфіденційності та етичних аспектів використання технологій розпізнавання облич у різних сферах, включаючи безпеку, медицину та соціальні системи. Надається огляд сучасних досягнень у сфері ідентифікації облич та вказано на перспективи подальших досліджень у цьому напрямку. Робота може бути корисною для фахівців у галузі комп'ютерного зору, машинного навчання, а також для тих, хто цікавиться етичними аспектами застосування технологій розпізнавання облич.

Ключові слова: розпізнавання обличь, аналіз зображень, алгоритми ідентифікації.

Zinovieva O. G. Analysis of face identification technologies in the image

The article is devoted to research and analysis of face identification methods in images. Current approaches and technologies used for recognizing persons in large volumes of data are considered. In particular, the methods of machine learning, neural networks and computer vision are investigated, which are used to improve the accuracy of identification and recognition of persons in images.

Image face recognition is a computer vision task aimed at recognizing and classifying faces of people or other objects in photos or videos. The task of face recognition has many applications in fields such as biometrics, organization of video conferences, machine vision systems in robotics, intelligent security and access control systems, etc. The task of detecting a face in an image is often the first step in the process of solving a higher-level problem – recognition of a face, details of a face or facial expressions. In addition, information about the presence and number of people in the image can be useful in systems of automatic accounting of the number of visitors; access control systems in institutions, airports and subways; automatic accident prevention systems; intelligent human-computer interfaces; in photography for automatic focusing on a person's face, as well as for stabilizing the image of a face in order to facilitate the recognition of emotions; to expand the area of stereo vision during the creation of 3D display systems. Analysis of face identification methods on images plays an important role in the development of computer vision and pattern recognition technologies.

The article highlights the main challenges in this direction, such as ensuring resistance to changes in lighting, facial position and external factors. Issues of privacy and ethical aspects of the use of facial recognition technologies in various fields, including security, medicine and social systems, are also explored. An overview of modern achievements in the field of face identification is provided and the prospects for further research in this direction are indicated. The work may be useful for specialists in the field of computer vision, machine learning, as well as for those interested in the ethical aspects of the use of face recognition technologies

Key words: face recognition, image analysis, identification algorithms.

Постановка проблеми. Зважаючи на частіші випадки тероризму, в сучасному світі особливо гостро постає проблема забезпечення безпеки людини в громадському транспорті, офісних приміщеннях, громадських місцях і т.д., тому завдання ідентифікації людини на зображенні, особливо по обличчю, набуває найімовірно важливого значення. Вирішення завдання ідентифікації осіб передбачає етапи отримання зображення, попередньої обробки, пошуку та виявлення осіб та подальшої ідентифікації особи з урахуванням знайдених характеристик.

Проблема ідентифікації осіб має ключову роль у системах відеоспостереження (наприклад, CCTV), а також сучасних охоронних комплексах. Також завдання ідентифікації осіб може застосовуватися в різних взаємодіях людини з машиною, наприклад, сучасні соціальні мережі виконують пошук за завантаженими зображеннями в соціальну мережу, ідентифікуючи знайдених людей, і пропонуючи їм асоціювати фотографію з їх акаунтом в мережі. Ще одним варіантом застосування завдання ідентифікації осіб є область «доповненої реальності», що набрала популярності останнім часом.

На сьогоднішній день сфера застосування алгоритмів виявлення осіб динамічно розвивається. Дані алгоритми знаходять застосування в різних системах, що вбудовуються, а умови застосування даних систем обумовлюють суттєві відмінності в якості зображень. Вимоги обробки в режимі реального часу унеможливають пост-обробку зображень або залучення оператора для, тому важливо розробляти стійкі до дефектів зображень алгоритми, що мають обчислювальну ефективність.

Таким чином, актуальність роботи полягає в тому, що завдання ідентифікації осіб є одним з найважливіших напрямків розвитку різних сучасних систем, таких як системи безпеки, системи машинно-людської взаємодії і т.д.

Аналіз останніх досліджень і публікацій. Проблема розпізнавання осіб розглядалася ще на ранніх стадіях комп'ютерного зору. Ряд компаній протягом понад 40 років активно розробляють автоматизовані, а зараз і автоматичні системи розпізнавання людських осіб: Smith & Wesson (система ASID – Automated Suspect Identification System); ImageWare (система FaceID); Imagis, Epic Solutions, Spillman, Miros (система Trueface); Vissage Technology (система Vissage Gallery); Visionics (система FaceIt). Багато методів ідентифікації осіб засновані на принципі, який вперше був застосований для розпізнавання задач, викладеному в [1].

Проблема формалізації та автоматизації процесу розпізнавання людських осіб була порушена ще на ранніх стадіях розвитку систем розпізнавання образів і залишається актуальною досі. Однак протягом останніх 10 років кількість наукових досліджень і публікацій з цієї тематики зростає в кілька разів, що свідчить про зростання актуальності цієї проблеми. За останні роки було запропоновано безліч різних алгоритмів обробки, локалізації та розпізнавання осіб такі, як нейронні мережі (Fe-rand, Rowley), власні особи (Turk, Pentland), ланцюги Маркова (Slimane) тощо.

Виклад основного матеріалу. Ключовим кроком на шляху до ідентифікації особи людини на зображенні є виявлення обличчя. Існує безліч різних технологій та алгоритмів виявлення осіб на зображеннях. Розглянемо найзначніші їх.

OpenCV (Open Source Computer Vision Library) – це відкрита бібліотека комп'ютерного зору та машинного навчання, яка надає різноманітні інструменти для обробки зображень та відео. Ця бібліотека має реалізації для різних мов програмування, таких як C++, Python, Java, і інших.

Крім кросплатформенності та підтримки багатьох мов програмування, які дозволяють використовувати програми на різних системах, бібліотека OpenCV дуже ефективна (в порівнянні з іншими схожими бібліотеками) з точки зору обчислень, так як майже всі функції та оператори в ній векторизовані

Основні модулі бібліотеки можна віднести до 4 груп (розділів):

- Модулі `core`, `highgui`, що реалізують базову функціональність (базові структури, математичні функції, генератори випадкових чисел, лінійна алгебра, швидке перетворення Фур'є, введення/виведення зображень і відео, введення/виведення у форматах XML, YAML та ін.).

- Модулі `imgproc`, `features2d` для обробки зображень (фільтрація, геометричні перетворення, перетворення кольорних просторів, сегментація, виявлення особливих точок та ребер, контурний аналіз та ін.).

- Модулі `video`, `objdetect`, `calib3d` (калібрування камери, аналіз руху та відстеження об'єктів, обчислення положення у просторі, побудова карти глибини, детектування об'єктів, оптичний потік).

- Модуль `ml`, що реалізує алгоритми машинного навчання (метод найближчих сусідів, наївний класифікатор байєсу, дерева рішень, бустинг, градієнтний бустинг дерев рішень, випадковий ліс, машина опорних векторів, нейронні мережі та ін.).

На особливу увагу заслуговує підтримка роботи з xml-файлами. Результати проміжних обчислень можна зберігати в xml-файлах, а потім їх прочитати, наприклад, в іншій програмі, що полегшує розробку алгоритму частинами, роботу над алгоритмом цілою командою.

Також варто зауважити, що бібліотека OpenCV реалізує, як правило, лише базові операції, які використовуються у комп'ютерному зорі. Таким чином, її можна розглядати як загалом низькорівневу бібліотеку комп'ютерного зору. Для вирішення серйозних завдань необхідно на основі наданих бібліотекою модулів створювати складніші програми, що реалізують необхідні завдання комп'ютерного зору.

Для вирішення завдання виявлення образів у бібліотеці OpenCV реалізовано алгоритм, що ґрунтується на ознаках Хаара. Ознаки Хаара – ознаки цифрового зображення, які використовуються для розпізнавання образів. Своєю назвою вони завдячують інтуїтивним подібності до вейвлетів Хаара. Ознаки Хаара використовувалися у першому детекторі осіб, що у реальному часі.

Історично склалося так, що алгоритми, що працюють тільки з інтенсивністю зображення (наприклад, значення RGB у кожному пікселі), мають велику обчислювальну складність. У роботі Папагеоргіу була розглянута робота з безліччю ознак, заснованих на вейвлетах Хаара. Віола та Джонс адаптували ідею використання вейвлетів Хаара та розробили те, що було названо ознаками Хаара. Ознака Хаара складається із суміжних прямокутних областей. Вони позиціонуються на зображенні, далі підсумовуються інтенсивності пікселів в областях, після чого обчислюється різниця між сумами. Ця різниця і буде значенням певної ознаки, певного розміру, певним чином позиціонованого на зображенні.

Наприклад, можна розглянути базу даних з людськими особами. Спільним для всіх зображень є те, що область в районі очей темніша, ніж область в районі щік. Отже загальною ознакою Хаара для осіб є 2 суміжні прямокутні регіони, що лежать на очах і щоках.

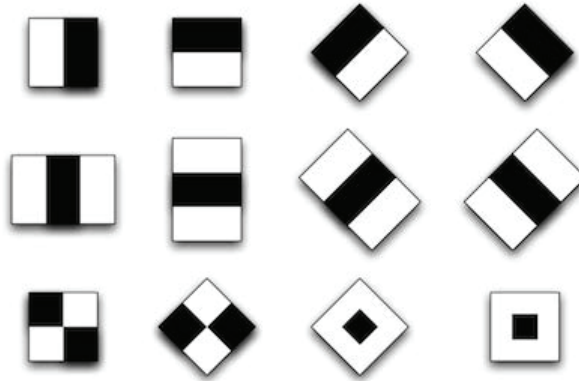


Рис. 1. Приклад ознак Хаара

На етапі виявлення у методі Віоли – Джонса вікно встановленого розміру рухається зображення, і кожній області зображення, над якою проходить вікно, розраховується ознака Хаара. Наявність або відсутність предмета у вікні визначається різницею між значенням ознаки та порогом, що навчається. Оскільки ознаки Хаара мало підходять для навчання або класифікації (якість трохи вища ніж у випадкової нормально розподіленої величини), для опису об'єкта з достатньою точністю необхідно більше ознак. Тому в методі Віоли-Джонса ознаки Хаара організовані в каскадний класифікатор.

Ключовою особливістю ознак Хаара є найбільша, проти іншими ознаками, швидкість. При використанні інтегрального представлення зображення ознаки Хаара можуть обчислюватися за постійний час (приблизно 60 процесорних інструкцій на ознаку з двох областей).

Бібліотека dlib

Dlib – це універсальна програмна бібліотека, написана мовою програмування C++. На його дизайн значною мірою впливають ідеї від проектування за контрактом та розроблення програмного забезпечення на основі компонентів. Таким чином, це насамперед набір незалежних програмних компонентів. Це програмне забезпечення з відкритим кодом, випущене під ліцензією Software Boost.

Основні функції та особливості бібліотеки dlib включають:

- розпізнавання облич: Dlib має реалізації для різних методів розпізнавання облич, включаючи багатопланові перцептрони та зглиблені згорткові мережі;
- виявлення об'єктів: бібліотека має ефективні алгоритми для виявлення об'єктів на зображеннях, що може бути використано для різних завдань;
- трекінг об'єктів: Dlib надає засоби для трекінгу об'єктів в реальному часі на відеопослідовностях;
- оптимізовані імплементації алгоритмів машинного навчання;

- обробка обличчя та обличчя-ключі: Dlib дозволяє визначати ключові точки на обличчі, використовуючи свої алгоритми, що може бути важливим для відомих задач, таких як розпізнавання емоцій;
- робота із зображеннями та відео: бібліотека має засоби для роботи зі зображеннями та відео, включаючи обробку та аналіз кадрів;
- зручний інтерфейс для Python.

Dlib є популярною серед розробників завдяки своїм потужним можливостям та оптимізованим реалізаціям. Бібліотека широко використовується в дослідженнях, розробці програм та прикладних рішеннях у сферах комп'ютерного зору та машинного навчання

Для виявлення осіб у бібліотеці використовується алгоритм, що базується на гістограмах спрямованих градієнтів, у поєднанні з лінійним класифікатором та методом ковзних вікон.

Гістограма спрямованих градієнтів (англ. Histogram of Oriented Gradients, HOG) – дескриптори особливих точок, які використовуються в комп'ютерному зорі та обробці зображень з метою розпізнавання об'єктів. Ця техніка заснована на підрахунку кількості напрямків градієнта у локальних областях зображення. Цей метод схожий на гістограми напрямку краю, дескриптори SIFT і контексти форми, але відрізняється тим, що обчислюється на щільній сітці рівномірно розподілених осередків і використовує нормалізацію локального контрасту, що перебивається, для збільшення точності [4].

Навніт Далал і Білл Тріггс, дослідники INRIA, вперше описали гістограму спрямованих градієнтів у своїй роботі на CVPR у червні 2005 року. У цій роботі вони використовували алгоритм для знаходження пішоходів на статичних зображеннях, хоча згодом розширили сферу застосування до знаходження людей на відео, а також різних тварин та машин на статичних зображеннях.

Основною ідеєю алгоритму є припущення, що зовнішній вигляд та форма об'єкта на ділянці зображення можуть бути описані розподілом градієнтів інтенсивності або напрямом країв. Реалізація цих дескрипторів може бути здійснена шляхом поділу зображення на маленькі зв'язкові області, іменовані осередками, і розрахунком для кожного осередку гістограми напрямків градієнтів або напрямків країв для пікселів, що знаходяться всередині осередку. Комбінація цих гістограм і є дескриптором. Для збільшення точності локальні гістограми зазнають нормалізації за контрастом. З цією метою обчислюється міра інтенсивності на більшому фрагменті зображення, який називається блоком, та отримане значення використовується для нормалізації. Нормалізовані дескриптори мають кращу інваріантність по відношенню до освітлення.

Дескриптор HOG має кілька переваг над іншими дескрипторами. Оскільки HOG працює локально, метод підтримує інваріантність геометричних і фотометричних перетворень, крім орієнтації об'єкта. Такі зміни з'являються лише у великих фрагментах зображення. Більш того, як виявили Далал і Тріггс, грубе розбиття простору, точне обчислення напрямків та сильна локальна фотометрична нормалізація дозволяють ігнорувати рух пішоходів, якщо вони підтримують вертикальне положення тіла. Дескриптор HOG, таким чином, є добрим засобом знаходження людей на зображеннях.

Першим кроком обчислень у багатьох детекторах особливих точок є нормалізація кольору та гамма-корекція. Далал і Тріггс встановили, що для дескриптора HOG цей крок можна опустити, оскільки подальша нормалізація дасть той самий результат. Тому на першому етапі розраховуються значення градієнтів. Найпоширенішим

методом є застосування одновимірної маски, що диференціює, в горизонтальному і/або вертикальному напрямку. Цей метод вимагає фільтрації колірної або яскравості за допомогою наступних фільтруючих ядер: $[-1,0,1]$ і $[-1,0,1]^T$.

Далал і Тріггс використовували більш складні маски, такі як Собел 3×3 або діагональні маски, але ці маски показали нижчу продуктивність для цього завдання. Вони також експериментували з розмиттям по Гауссу перед застосуванням маски, що диференціює, але також виявили, що пропуск цього кроку збільшує швидкість без помітної втрати якості.

На наступному етапі обчислюються гістограми осередків. Кожен піксель у осередку бере участь у зваженому голосуванні для каналів гістограми напрямів, заснованому значенні градієнтів. Осередки можуть бути прямокутної або круглої форми, канали гістограми рівномірно розподіляються від 0 до 180 або від 0 до 360 градусів, залежно від того, обчислюється «знаковий» або «беззнаковий градієнт». Далал і Тріггс виявили, що беззнаковий градієнт разом із дев'ятьма каналами гістограми дає найкращі результати при розпізнаванні людей. При розподілі ваги в голосуванні вага пікселя може задаватися або абсолютним значенням градієнта, або деякою функцією від нього; у реальних тестах абсолютне значення градієнта дає найкращі результати. Іншими можливими варіантами можуть бути квадратний корінь, квадрат або абсолютне урізане значення градієнта.

Для врахування яскравості і контрастності градієнти слід локально нормувати, навіщо осередки необхідно згрупувати у великі зв'язкові блоки. Дескриптор HOG, таким чином, є вектором компонента нормованих гістограм осередків з усіх областей блоку. Кінцевим кроком розпізнавання об'єктів з використанням HOG є класифікація дескрипторів за допомогою системи навчання з учителем.

Порівняння бібліотек виявлення осіб

У таблиці 1 наведено результати порівняння технологій виявлення облич OpenCV (вейвлети Хаара), dlib (гістограми спрямованих градієнтів, HOG), а також примітивний алгоритм SIFT (Scale-invariant feature transform). Як набір зображень були використані бази даних пішоходів Массачусетського технологічного інституту (MIT), а також база даних людей Національного дослідницького інституту Франції (INRIA). База даних INRIA відрізняється тим, що в ній зображені люди у різних позах та ракурсах.

Таблиця 1

Результати порівняння технологій виявлення осіб

Алгоритм	Набір зображень	Частка пропущених зображень	Частка помилок першого роду
HOG	MIT	~ 0	10-4
HOG	INRIA	0,1	10-4
Вейвлети Хаара	MIT	0,01	10-4
Вейвлети Хаара	INRIA	0,3	10-4
SIFT	MIT	0,1	10-4
SIFT	INRIA	0,5	10-4

Висновки. З результатів проведеного дослідження випливає, що на вибірці людей, звернених в камеру безпосередньо, обидва алгоритми показують приблизно однакові результати, проте на зображеннях людей під різним ракурсом, алгоритм, що використовує метод ознак Хаара, значно поступається алгоритму,

заснованому на гістограмах градієнтів. У свою чергу, примітивний алгоритм SIFT явно поступається обом алгоритмам, що доводить їхню актуальність.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ:

1. Edwards G. J. Cootes T. F. Taylor C. J. Face recognition using active appearance models. *Computer Vision*. Volume 1407 of the series Lecture Notes in Computer Science, 2006, p. 581–595.
2. S. Sankaranarayanan, A. Alavi, C. Castillo, R. Chellappa, Triplet Probabilistic Embedding for Face Verification and Clustering. *2016 IEEE 8th International Conference on Biometrics Theory, Applications and Systems (BTAS)*. 2016, p. 1–8.
3. T. Zhang, Lu Bao-Liang “Selecting optimal orientations of Gabor wavelet filters for facial image analysis”, 2010, p. 218–227.
4. Довбиш, А.С. Основи теорії розпізнавання образів: навч. посіб. Суми: Сумський державний університет, 2015. Ч. 1. 109 с.
5. Вовк С.М., Гнатушенко В.В., Бондаренко М.В. Методи обробки зображень та комп'ютерний зір: Навчальний посібник. Дніпро: ЛІРА, 2016. 148 с.
6. Лисенко А. М. Застосування біометричних систем для ідентифікації особи. Вісник Київського нац. ун.-ту ім. Т.Шевченка, Юридичні науки, 2004, №60/62, с. 87–91.

REFERENCES:

1. Edwards G. J. Cootes T. F. Taylor C. J. (2006). *Face recognition using active appearance models*. *Computer Vision*. Volume 1407 of the series Lecture Notes in Computer Science, p. 581–595.
2. Sankaranarayanan, S., Alavi, Castillo, C & Chellappa, R. (2016) *Triplet Probabilistic Embedding for Face Verification and Clustering*. *2016 IEEE 8th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, p. 1–8.
3. T. Zhang & Lu Bao-Liang (2010). *Selecting optimal orientations of Gabor wavelet filters for facial image analysis*, p. 218–227.
4. Dovbysh, A.S. (2015). *Basics of pattern recognition theory*. Sumy: Sumy State University, Vol. 1.
5. Vovk, S. M., Hnatushenko, V.V. & Bondarenko, M.V. (2016)/ *Image processing methods and computer vision*. Dnipro: LIRA.
6. Lysenko A. M. (2004)/ *Application of biometric systems for personal identification*/ *Visnyk Kyivskoho nats. un.-tu im. T.Shevchenka, Yurydychni nauky*, №60/62, pp. 87–91.