



ТЕХНОЛОГІЯ РОЗПІЗНАННЯ ОСОБИСТОСТІ У ВІДЕОПОТОКУ В РЕЖИМІ РЕАЛЬНОГО ЧАСУ

Супрун О.О., доцент, кафедра МІРЕС, ХНУРЕ
Кошель В.О., студент, кафедра МІРЕС, ХНУРЕ

Завдання розпізнавання осіб є одним із перших практичних завдань, яке послужило стимулом для розвитку теорії розпізнавання об'єктів. Останнім часом у різних сферах діяльності, пов'язаних із обробкою інформації, зростає потреба швидкої та правильної ідентифікації особистості людини у відеопотоці в режимі реального часу.

У загальному випадку завдання ідентифікації особи у відеопотоці може бути поділено на кілька етапів: пошук області особи на зображенні; стеження за положенням виявленої особи на послідовності кадрів (трекінг); попередня обробка зображення особи з метою компенсації умов зйомки; порівняння зображення обличчя із зображеннями в базі зображень. Трекінг є не обов'язковим етапом, але дозволяє співвідносити зображення осіб людей різних кадрах і робити розпізнавання за найкращим зображенням з послідовності кадрів, побічно підвищуючи точність ідентифікації.

Завдання розпізнавання особистості є більш складним, і на даний момент існуючі алгоритми наближаються до точності розпізнавання осіб людиною тільки в заданих умовах, або вимагають обчислювальні ресурси, недоступні при обробці відеопотоку в режимі реального часу.

Число помилок ідентифікації особистості у всіх сучасних систем розпізнавання осіб у режимі реального часу має явну залежність від якості зображень осіб, механічних перешкод, ракурсу зйомки особи, що ідентифікується, умов освітлення.

Для вирішення цих проблем існує два шляхи: розширення бази еталонних зображень зображеннями з різними умовами зйомки (що не завжди можливо зробити), а також моделювання умов зйомки на еталонних зображеннях за їх порівняння з аналізованими.

Існуючі швидкі 2D методи та алгоритми досягають гарної точності тільки при дотриманні жорстких вимог щодо ракурсу зображень обличчя та умов зйомки, що обумовлено механізмами кластеризації та машинного навчання.

Попередня обробка може проводитися за кількома критеріями, такими як контрастність, різкість, освітлення, ракурс. Технологію попередньої обробки можна представити у вигляді комплексного алгоритму для завдання обробки відеопотоку, в тому числі і в режимі реального часу, який повинен враховувати всі критерії і мати достатню швидкодію.

Проблема попередньої обробки зображень найчастіше розглядається як набір приватних завдань щодо покращення якості, компенсації освітлення та ракурсу зйомки, але не як складна система взаємозалежних алгоритмів.

Необхідність оптимізації всіх кроків обробки та аналізу зображень осіб під конкретне завдання розпізнавання, чи то фотографії чи відеопотік,



створюють передумови для створення комплексної системи алгоритмів. Для отримання докладнішої 3D моделі обличчя необхідно розширити масив до певної межі $Q \leq N$, де N – кількість точок, що описують вихідні 3D моделі. Розбиття отриманої моделі обличчя на трикутники. В результаті цієї операції буде отримано набір трикутників, що описує область обличчя на зображенні. Дроблення трикутників за вибраним критерієм (тесселяція), що дозволяє знизити розміри трикутників, що використовуються, і наблизити трикутну сітку до рівномірної. Як критерій розбиття можна брати максимальну площу трикутника або максимальну довжину сторони.

Кожен із трикутників за допомогою операцій обертання щодо заданого центру та афінного перетворення зміщується та деформується відповідно до заданої величини кута. В результаті виконання алгоритмів з одного зображення обличчя вдається отримати набір зображень, який можна використовувати в алгоритмах розпізнавання особистості.

Даний метод є простим з обчислювальної точки зору і для систем розпізнавання може застосовуватися за умовчанням, незважаючи на низьку точність через необхідність використання детектора Віоли-Джонса для виділення областей осіб. Для зниження обчислювальної складності детектора осіб пропонується попередній аналіз кадрів відео на наявність областей, що потенційно містять особи, методом віднімання фону і виявлення руху background subtraction.

До запуску детектора додається операція обчислення різниці між кадрами: $f_T = f_i - f_0$, $f_M = f_i - f_{i-1}$. Обчислюються бінарні зображення f_T і f_M з порогом, що залежить від якості відеопотоку. Далі ковзним вікном зі сторонами, рівними половині боку мінімального вікна детектора, проводиться оцінка сум білих пікселів. Ті області зображення, у яких сума білих пікселів перевищує обраний поріг, відзначаються як відповідні для подальшого аналізу. У цих областях перебувають рухомі об'єкти, зокрема особи людей.

Такий підхід дозволяє уникнути обчислення вейвлетів Хаара для всього кадру, значно знижуючи обчислювальну складність трекінгу, заснованого на детекторі Віоли-Джонса.

Для перевірки ефективності використання алгоритму автоматичного синтезування 3D моделей осіб під час формування набору зображень осіб проведено його тестування у складі системи розпізнавання, що була реалізована під час дослідження. Система заснована на методах обчислення біометричних векторів стандартним відхиленням (STD), зниження розмірності векторів Principal Component Analysis (PCA) та класифікатори k Nearest Neighbors (kNN). Отримані вектори ознак порівнювали за метрикою Евкліда.

При порівнянні результатів розпізнавання як оцінка ефективності алгоритму часто вибирається характеристика Equal Error Rate (EER). EER – рівний рівень помилок FAR і FRR, де FAR – можливість помилкового виявлення, тобто. ймовірність те, що система помилково визнає справжність користувача, не зареєстрованого у системі, а FRR – ймовірність пропуску мети,



тобто. ймовірність того, що система не визнає справжність зареєстрованого користувача. Однак, така оцінка не завжди відображає реальну ефективність, так як вона залежить не тільки від мінімальних або оптимальних значень оцінок FAR і FRR, але і від динаміки їх зміни, сфери застосування алгоритмів. У системах, що контролюють вхід людей на об'єкти обмеженого допуску, можливість пропуску мети не така критична характеристика, як можливість помилкового виявлення. У проведених експериментах оцінкою ефективності системи ідентифікації особи прийнято рівень FRR при заданому рівні FAR=1% вибірки. У разі використання системи розпізнавання для пошуку в кадрі конкретної особи, що розшукується, помилка FRR стає більш значущою, і в таких завданнях є сенс мінімізувати саме її.

Отримані результати об'єднані у вигляді таблиці рівня помилки FRR при фіксованому рівні FAR, а також у вигляді графіка, що показує співвідношення рівнів помилок FAR і FRR для різних способів формування набору вхідних зображень для системи розпізнавання особистості.

Таблиця 1 – Результати тестування системи розпізнавання особистості з урахуванням зображень обличчя

База вихідних зображень	Власна	Caltech faces	FERET
Спосіб формування бази	FRR, % (при FAR=1%)		
Одне зображення на людину	8,07	5,19	3,56
15 зображень на людину	4,82	2,35	0,88
15 згенерованих алгоритмом зображень на людину	2,18	3,46	3,17

Як можна бачити з таблиці 1, результати тестування вибраної метрики підтверджують ефективність алгоритму розширення набору вихідних зображень: на власній базі точність розпізнавання підвищилася в порівнянні з набором зображень з самої бази. На базах Caltech Faces та FERET тестування показало, що набір синтезованих зображень підвищує точність розпізнавання порівняно з базою одиночних зображень.

Це дозволяє стверджувати, що використання синтетичних зображень в різних ракурсах, отриманих за допомогою запропонованої модифікації DFM, підвищує точність розпізнавання особи за зображенням обличчя на 0,5-4% в порівнянні з обмеженим набором зображень.

Список літератури

1. Viola, P., & Jones, M.J. (2001). Rapid Object Detection using a Boosted Cascade of Simple Features. Computer Vision and Pattern Recognition (CVPR 2001).
2. Viola, P., & Jones, M.J. (2004). Robust real-time face detection. International Journal of Computer Vision, 57(2), 137-154.
3. Гонсалес, Р., & Вудс, Р. (2005). Цифрова обробка зображення. ISBN 5-94836- 028-8.
4. Sochman, Y., & Matas, Jiří. (2010). AdaBoost. Center for Machine Perception, Czech Technical University, Prague.