

СИСТЕМИ РОЗПІЗНАВАННЯ ОСІБ ЗА ДОПОМОГОЮ 3D

Системи розпізнавання осіб застосовуються для електронної ідентифікації людей. Чим більша відстань від об'єкта до сканера, тим просунуте рішення потрібно для точного результату. Якщо камера знаходиться далеко, системі доведеться шукати обличчя людей, відокремлювати їх від натовпу, нахилити, розтягувати та спотворювати зображення, щоб виділити риси обличчя та вирівняти їх для точного виміру.

Сучасні системи з технологією розпізнавання облич працюють з великої відстані і навіть у натовпі можуть швидко обчислити потрібну людину. При спостереженні за тривимірним світом за допомогою 2D-камери втрачається інформація, необхідна для відеоаналізу. Причина невдач 2D-аналітики в тому, що світ тривимірний, а зображення двовимірне. Недолік інформації заповнюється гіпотезами та припущеннями, які, як підтверджують відомі оптичні ілюзії, не завжди виправдовуються. Опис тривимірних об'єктів у термінах двовимірних проєкцій складний і громіздкий, тому двовимірна відеоаналітика найчастіше справляється тільки із відносно простими завданнями, наприклад з ідентифікацією об'єктів.

Природний крок – перехід до тривимірного відеоспостереження. Відомо кілька способів отримання тривимірного зображення. Один із них, лазерне сканування, займає досить тривалий час – десятки хвилин, тому відеоспостереження за людьми не підходить. Інший метод – структуроване підсвічування у вигляді горизонтальних ліній та аналіз викривлення цих ліній (як у приставці Kinect). Недоліки – мала дальність та неможливість роботи на сонці. Третій спосіб заснований на факті кінцівки швидкості світла - вимірюється час відбиття світла від об'єктів на різній відстані. Цей перспективний метод зараз активно розвивається, але він має серйозні обмеження, передусім пов'язані з невеликою відстанню до об'єкта, який має висвітлюватися активним підсвічуванням.

Джерела проблеми розпізнавання осіб – неправильний ракурс (при ракурсі 15% жоден з існуючих алгоритмів не здатний розпізнати обличчя), міміка, макіяж, вуса та борода, окуляри. Тривимірний машинний зір допомагає вирішити багато перерахованих проблем. Для відеоаналітики використовують 3D-моделі об'єктів реального світу. Знання абсолютних координат та розмірів об'єктів спрощує їх сегментацію та виділення, дає можливість, наприклад, відокремити об'єкт від фону або розділити два об'єкти у кадрі. Ефективніше вирішується завдання ідентифікації об'єктів та підвищується достовірність розпізнавання осіб (по 3D-моделі особи) у різних ракурсах (див. рис. 6). За даними тривимірної траєкторії можна зробити висновки щодо поведінки об'єкта.

3D-система дозволяє побудувати об'ємну модель обличчя та розгорнути це синтетичне зображення для точного розпізнавання з використанням бази 2D-осіб. З удосконаленням аналітичних алгоритмів та збільшенням обчислювальних потужностей процесорів, які застосовуються в серверах та камерах відеоспостереження, можливості таких інструментів значно розширюються.

Система розпізнавання осіб, як будь-яка технологія, має похибки. Складність полягає у перетворенні біологічних характеристик (відбитків пальців, рис обличчя) в цифрові дані (математичну модель), на цьому етапі навіть мінімальний рух людини може призвести до невірних розрахунків.

Точність роботи системи можна вимірювати швидкістю розпізнавання (RR, Recognition Rate) - це загальна кількість правильно ідентифікованих зображень осіб людей, поділена на загальну кількість тестових зображень. Наприклад, 9 осіб із 10 розпізнаються вашою системою – ваш RR дорівнює 90%.

З іншої сторони, точність системи пов'язані з алгоритмом. Він виявляє обличчя і порівнює його з певним набором даних, який містить безліч зображень однієї і тієї ж людини або велику кількість зображень різних людей. Продуктивність алгоритму пов'язана з кількістю даних, задіяних під час навчання, для отримання точних результатів ви повинні навчати алгоритм на тисячах різних зображень. У якості критерія правильного розпізнавання можна вибрати точність суміщення отриманого зображення з еталонним для побудови 3D-моделі особи

$$\sigma_p^2 = \frac{\left(\frac{2}{q_1} + \frac{2}{q_2} + \frac{1}{q_1 q_2} \right)}{4k_{12}^2 \omega^2} = \frac{2q_1 + 2q_2 + 1}{4k_{12}^2 q_1 q_2 \omega^2}$$

де q_1, q_2 – відношення сигнал-шум на зображенні і еталоні відповідно, k_{12} – коефіцієнт кореляції спектрів зображень об'єкту на зображенні і еталоні, ω^2 - другий момент взаємного просторового спектру зображення і еталону.