

Підвищення якості ідентифікації людини за зображенням обличчя на основі використання лінійних підпросторів

Кирило В. Муригін

Відділ розпізнавання зорових образів
Інститут проблем штучного інтелекту, Донецьк
kir@iaai.donetsk.ua

Абстракт

У статті представлені експериментальні дослідження методу розпізнавання людини, що базується на еластичному зіставленні графів зображень облич. Зображення облич представляються у виді графів облич [1, 2, 3, 4] – сукупності точок і з'єднуючих їх ребер. Кожній точці (вузлу) графа обличчя відповідає вектор ознак, що характеризує особливості зображення в даній точці. Для розпізнавання використовується кореляційний метод порівняння графів зображень облич.

Метою досліджень, описаних у даній статті, є пошук більш зручного для рішення задачі розпізнавання базису в просторі ознак, що утворюються на основі згорток зображень облич із функціями Габора різних масштабів і кутів повороту [1, 2, 3, 4]. Передбачається, що базис, який потрібно знайти, є підпростором вихідного простору ознак.

Представлення зображень облич

Кожна точка довільного зображення характеризується деяким набором властивостей, що відбивають особливості зображення навколо неї. Ідеальним описом властивостей точки, мабуть, є яскравості сусідніх пікселів – тобто саме зображення поблизу точки, однак таке представлення найчастіше сильно надлишкове і не раціональне. Тому для формування ознак точки часто використовують коефіцієнти розкладання області зображення в її околиці по деякому базису. Таким чином, кожній точці вихідного зображення ставляться у відповідність вектора коефіцієнтів – вектора ознак точки. У дійсній роботі для представлення таких локальних властивостей використовуються коефіцієнти згорток зображення з набором функцій Габора (5 масштабів і 8 кутів повороту [1, 2, 3, 4]). Переваги використання як ознаки для розпізнавання облич векторів згорток зображення й функцій Габора різних масштабів і кутів повороту описані в роботах [5, 6].

Для порівняння векторів ознак, отриманих на основі згорток зображення з набором функцій Габора,

використовується міра у виді косинуса кута між ними:

$$S(J_1, J_2) = \frac{(J_1, J_2)}{|J_1||J_2|}. \quad (1)$$

Така функція порівняння інваріантна до модулів векторів ознак і в сполученні з властивістю незалежності згортки функцій Габора щодо постійного зсуву зображення по шкалі яскравості робить результат порівняння незалежним від будь-яких лінійних перетворень вихідних зображень, на основі яких вектора ознак були отримані. А це означає інваріантність результату порівняння стосовно рівномірних змін яскравості й контрастності зображень.

Згідно методу еластичного зіставлення графів, зображення облич трансформуються в граф обличчя, що містить їхні властивості. Конфігурація графа визначає геометричні особливості обличчя, вектора властивостей у вузлах – локальні властивості зображення. У дійсній роботі використовується конфігурація графа обличчя у виді прямокутної сітки, при цьому перша зверху лінія вузлів графа відповідає лінії брів; друга – лінії очей; четверта – лінії, що проходить через кінчик носа; шоста – лінії губ (див. рис. 1).

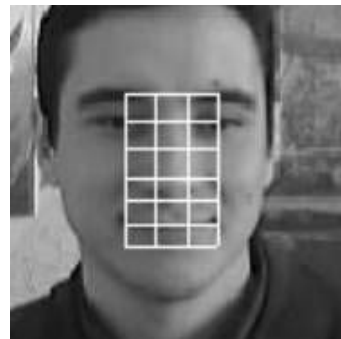


Рис. 1. Конфігурація графа обличчя, що використовується в роботі

Як видно з рис. 1, граф розташовується в центральній частині обличчя, що дозволяє виключити вплив на формування ознак зачіски і навколишнього фону – найбільш варіабельних областей зображення обличчя.

Для визначення близькості зображень облич, порівнюються відповідні їм графи облич за допомогою міри:

$$S_G(G^1, G^2) = \frac{1}{N} \sum_{n=0}^{N-1} S(G_n^1, G_n^2),$$

де S_G – результат порівняння графів; N – число вузлів у графі; n – індекс, що визначає номер вузла; G_n^1, G_n^2 – вектора ознак, що відповідають n -му вузлу графа; $S(G_n^1, G_n^2)$ – функція порівняння векторів ознак (1).

Ухвалення рішення про приналежність даного зображення класу зображень облич даної людини здійснюється за допомогою порогу на результат порівняння графів зображення, що тестується і графа обличчя відомої людини.

При аналізі чергового зображення обличчя, конфігурація графа підбудовується під його пропорції для того, щоб домогтися відповідності між порівнюваними точками зображень облич.

Дослідження простору ознак

Вектори ознак, що використовуються для розпізнавання, мають багато надлишкової інформації. У них містяться дані, як про індивідуальні особливості обличчя, так і про умови, при яких зображення обличчя було отримане. При цьому зміни зображень облич, пов'язані з невеликими поворотами чи зміною напрямку освітлення, виявляються більш істотними, ніж відмінності між зображеннями облич різних людей, що були отримані в однакових умовах. Так у роботі [7] показано, що при використанні для ідентифікації методу власних облич, відсоток помилок розпізнавання значно знижується, якщо виключити з розгляду перші три компоненти, що у цьому випадку відповідають найбільш істотним відхиленням усередині класу зображень всіх облич. Ґрунтуючись на тім, що напрямки змін векторів ознак, пов'язані зі зміною напрямку висвітлення чи іншими факторами внутрішньокласового розкиду, відмінні від змін, що відповідають відмінностям між зображеннями облич різних людей, можна припустити, що в просторі ознак існує підпростір меншої розмірності, у якому класи зображень облич різних людей розділяються щонайкраще. Тому дослідження з виділення напрямків, що відповідають за розпізнавання, і зменшення внеску напрямків, що більше

відповідають за внутрішньокласовий розкид можуть допомогти у визначенні більш зручного для класифікації простору ознак, що підвищить якість роботи системи ідентифікації.

При порівнянні векторів ознак використовуються тільки їхні проекції на знайдений підпростір. Нехай T^m , $m = 1..M$, $M < N = 40$ – ортонормований базис отриманого підпростору розмірності M для однієї з точок графа обличчя, N – розмірність вихідного простору ознак у даному випадку дорівнює 40 (5 масштабів \times 8 кутів повороту). Тоді для результату порівняння двох векторів ознак, що відповідають даній точці, одержимо:

$$S = \frac{\sum_{k=1}^N \left\{ \sum_{m=1}^M \alpha_m T_k^m \right\} \left\{ \sum_{n=1}^M \beta_n T_k^n \right\}}{\left| \sum_{m=1}^M \alpha_m T_k^m \right| \left| \sum_{n=1}^M \beta_n T_k^n \right|},$$

де $J_k^1 = \sum_{m=1}^M \alpha_m T_k^m$ і $J_k^2 = \sum_{n=1}^M \beta_n T_k^n$ – відповідно розкладання першого і другого векторів ознак по базису T^m . Звідси, з огляду на ортонормованість T^m , одержимо:

$$S = \frac{\sum_{m=1}^M \alpha_m \beta_m}{\sqrt{\sum_{m=1}^M \alpha_m^2} \sqrt{\sum_{m=1}^M \beta_m^2}}.$$

Таким чином, міра подібності двох векторів може бути обчислена через коефіцієнти розкладання по базису знайденого підпростору з використанням функції порівняння у виді косинуса кута між векторами цих коефіцієнтів.

Для рішення задачі пошуку підпростору, у якому підвищується ефективність розпізнавання, зручно використовувати метод облич Фішера (в англійській літературі часто використовується термін Fisherfaces), заснований на використанні лінійного дискримінанта Фішера [7]. Цей метод використовувався раніше для векторів ознак, що представляють собою повне зображення обличчя людини, і показав непогані результати. Метод облич Фішера, на відміну від відомого підходу власних облич, що виділяє напрямки найкращого розкладання довільного зображення обличчя по базису власних векторів, характерних простору облич, визначає напрямки, що відповідають максимуму відносини між класового розкиду до внутрішньокласового:

$$W_{opt} = \arg \max \frac{|W^T S_B W|}{|W^T S_W W|} = [w_1, w_2, \dots, w_m],$$

між класового

Напрямки, які потрібно знайти, є рішеннями наступної задачі на власні значення:

$$S_B w_i = \lambda_i S_W w_i.$$

Для розрахунку матриць S_B і S_W використовувалася база даних Weizmann Institute of science, що містить 160 зображень облич 10 чоловік (16 зображень для кожної людини). Усі фотографії були отримані при різних напрямках освітлення. Розрахунки здійснювалися за наступними формулами:

$$S_W^{ij} = \sum_{k=1}^K \sum_{p=1}^P (J_{ip}^{nk} - M_i^{nk}) \cdot (J_{jp}^{nk} - M_j^{nk}),$$

$$S_B^{ij} = \sum_{k=1}^K P \cdot (M_i^{nk} - M_i^n) \cdot (M_j^{nk} - M_j^n).$$

Тут K – кількість класів (людей у базі даних); P – число векторів у кожному класі (зображень кожної людини); k – індекс класу; p – індекс вектора в класі; n – номер точки графа, для якої виконується розрахунок; J_{ip}^{nk} – (i -й) елемент вектора ознак; M_i^{nk} – (i -й) елемент середнього по (k -му) класу вектора ознак; M_i^n – (i -й) елемент середнього по усіх векторах ознак для (n -й) точки графа обличчя.

Для знаходження зворотної матриці S_W використовувався метод, заснований на LU декомпозиції.

Для розрахунку власних векторів і власних значень матриці $S_W^{-1} \cdot S_B$ був використаний алгоритм, що складається з двох етапів. На першому етапі виконується алгоритм приведення симетричної матриці до трьох діагонального виду. Потім виконується QL алгоритм розрахунку власних векторів і власних значень.

Після того як набори власних векторів для кожної точки графа обличчя були розраховані, були проведені дослідження з визначення залежності відсотка помилок розпізнавання від числа використовуваних власних векторів. Для цього була застосована та ж база даних зображень облич, що і при розрахунку матриць між класового й внутрішньокласового розкиду. Однак при розрахунку векторів ознак графа обличчя

розташовувався автоматично, за допомогою алгоритму, описаного в [8]. На основі аналізу функцій розподілу для результатів порівняння графів зображень облич однієї людини й облич різних людей були розраховані значення для помилок першого й другого роду. Отримані дані приведені нижче на рис. 2. Число використовуваних власних векторів при проведенні розрахунків було однаковим для кожної точки графа обличчя.

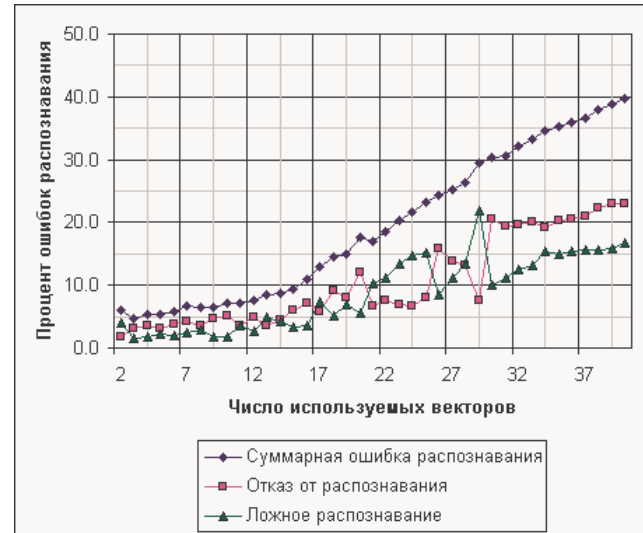


Рис. 2. Залежність відсотка помилок розпізнавання від числа використовуваних власних векторів, знайдених по методу Фішера

З графіків, показаних на рис. 2, видно, що зниження розмірності простору ознак за рахунок використання оптимальних проєкцій Фішера, значно зменшує відсоток помилок (з 40% до 7%). При цьому максимальна ефективність роботи системи розпізнавання досягається при використанні 2-7 мірного простору ознак замість, раніше використовуваних, 40 вимірів. Це дозволяє істотно скоротити розміри бази даних ознак і підвищити швидкість роботи алгоритму розпізнавання. Крім методу оптимальних проєкцій Фішера, також був досліджений його окремий випадок, коли матриця між класового розкиду є одиничною. У цьому випадку якість розпізнавання підвищується за рахунок зменшення внутрішньокласового розкиду. Таке наближення засноване на припущенні про те, що усередині всіх можливих класів, що відповідають обличчям різних людей, напрямки змін векторів ознак, пов'язані зі змінами освітлення й інших факторів, приблизно однакові. У цьому випадку задача на власні значення приймає вид:

$$w_i = \lambda_i S_W w_i, \text{ чи } S_W^{-1} w_i = \lambda_i w_i.$$

Отримане рішення було тестовано на тій же базі даних векторів ознак, що і базис, розрахований на основі лінійного дискримінанта Фішера. Графік залежності відсотка помилок розпізнавання від числа використовуваних напрямків простору ознак приведений на рис. 3 і рис. 4.



Рис. 3. Залежність відсотка сумарної помилки розпізнавання від числа використовуваних напрямків, знайдених на основі мінімізації внутрішньокласового розкиду



Рис. 4. Залежність відсотка сумарної помилки розпізнавання від числа використовуваних напрямків, знайдених на основі мінімізації внутрішньокласового розкиду (конкретизований діапазон 3-26)

З графіка видно, що при використанні більше 22 векторів, відсоток помилок розпізнавання різко збільшується з ростом числа векторів. Це пов'язано з різким збільшенням впливу факторів внутрішньокласового розкиду (наприклад, умов висвітлення). Оптимальним є використання 13 векторів, при цьому досягається мінімум відсотка сумарної помилки розпізнавання – приблизно 0.01%.

Висновки

Проведені дослідження простору ознак, отриманих у виді коефіцієнтів згорток зображень із набором функцій Габора різних масштабів і кутів повороту,

показали можливість зниження його розмірності при зниженні сумарної помилки розпізнавання. Найбільш істотне зменшення помилок розпізнавання показав метод знаходження базису підпростору, заснований на зменшенні внутрішньокласового розкиду в просторі ознак. Це говорить про подібність напрямків змін, що відбуваються в зображеннях обличчя різних людей при східних змінах у напрямку освітлення. Зменшення розмірності простору ознак дозволило покращити швидкодію методу розпізнавання та зменшити обсяг інформації, що зберігається в базі даних обличчя.

Література

- [1] M. Lades et al. Distortion invariant object recognition in the dynamic link architecture. *IEEE Trans. Comput.*, 42(3): 300-311,1993.
- [2] Wiskott L., Fellous J.M., Krueger N. and von der Malsburg C. Face Recognition and Gender Determination. // *Proc. of the Int. Workshop on Automatic Face- and Gesture-Recognition, Zuerich, 1995.* -p.92-97.
- [3] Wiskott L., Fellous J.M., Krüger N., von der Malsburg C. Face recognition by elastic bunch graph matching //*IEEE Trans. on Pattern Analysis and Machine Intel.* -1997. -Vol.19, №7. -P.775-779.
- [4] Wiskott L. Phantom faces for analysis //*Proceedings of Int. Conf. on Computer Analysis of Images and Patterns.* -1997. -Vol.1256. -P.480-487.
- [5] Мурыгин К.В. Исследование масштабной и угловой зависимости векторов признаков, полученных на основе сверток с функциями Габора //*Искусственный интеллект.* -2002. -№4. -С.499-502.
- [6] Мурыгин К.В. Оптимизация габоровских вейвлет для задачи распознавания человека по изображению лица //*Искусственный интеллект.* - 2003. -№4. -С.223-229.
- [7] Belhumeur P. N., Hespanha J. P. and Kriegman D. J. Eigenfaces vs Fisherfaces: Recognition Using Class Specific Linear Projection //*IEEE Transactions on Pattern Analysis and Machine Intelligence.* -1997. -Vol.19. -P.711-720.
- [8] Мурыгин К.В. Выявления ключевых точек на зображенні обличчя людини //*Матеріали VI Всеукраїнської Міжнародної конференції "Оброблення сигналів і зображень та розпізнавання образів".* - Київ. -2002. -С.123-126.