

ВИКОРИСТАННЯ "ВЛАСНИХ ОБЛИЧ" ДЛЯ РОЗПІЗНАВАННЯ ЛЮДСЬКИХ ОБЛИЧ

Левко О. В., Нюнькін К. М.

Інститут проблем штучного інтелекту

83050, м.Донецьк, проспект Богдана Хмельницького, 84

Тел.: (0622)926082, факс: (0622)926082

E-mail: kmn@iai.donetsk.ua

ABSTRACT

In this paper usage of a method of image decomposition on eigenfaces for human face recognition is considered. The experiments permitting are conducted to define limitations of this approach.

ВСТУП

Комп'ютерне розпізнавання людських облич привертає в останній час все більше уваги. З одного боку це пов'язано з численними задачами, в яких може застосовуватися ця техніка, з іншого - навалює зростаючі потужності сучасних комп'ютерів роблять можливим вирішення таких задач.

Всю безліч підходів до розпізнавання облич можна поділити на дві великі групи: методи, що використовують окремі деталі зображення, й алгоритми, що розглядають все зображення цілком. В роботі розглядається один з підходів другий групи - розклад зображення по "власним обличчам" або розклад Karhunen-Loeve [1.2], що вже досить давно використовується для стиску зображень. Поняття "власні обличчя" вводяться по аналогії з власними векторами і мають ту же властивість: їх можна вибрати як базис, по якому може бути розкладений довільний елемент простору облич.

1. ПРОСТІР ОБЛИЧ

Зображення в комп'ютері представляється матрицею окремих пікселів, яка має V рядків по N пікселів у кожному. У випадку сірого напівтонового зображення значення пікселів визначаються яскравістю відповідної точки зображення. Таке зображення можна уявити у вигляді вектору. Для цього необхідно з'єднати послідовно всі рядки зображення, отримавши один довгий рядок. Значення кожного пікселя зображення буде

відповідати одній компоненті вектору. Отриманий вектор належить простору усіх можливих зображень, що мають розмір $N \times V$ пікселів. Цей вектор має вимірність $N \times V$. Ясно, що працювати з векторами такої великої вимірності дуже складно.

При вирішенні задачі розпізнавання людських облич нас цікавлять не будь-які зображення, а тільки зображення облич. Всі обличчя схожі друг на друга: всі вони мають ніс, очі, губи й інші характерні деталі, розташовані на певних місцях. Тому зображення облич в просторі зображень компактно згруповані і утворюють порівняно невеликий кластер. Отже, повний простір зображень є надмірним для опису облич і можна побудувати спеціальний простір, що буде описувати обличчя оптимальним чином, маючи при цьому значно меншу вимірність, ніж вихідний простір зображень. Природно цей простір назвати простором облич [3].

Нехай $I=N \times V$ - кількість пікселів у зображенні та число компонентів вектору простору зображень; K - кількість зображень облич в навчальній вибірці; X - матриця вимірності $K \times I$, що визначається навчальною вибіркою. k -ий стовбчик матриці відповідає k -му зображенню обличчя навчальної множини.

Надмірність простору зображень є наслідком того, що кожний піксель обличчя корелює з іншими пікселями і коваріаційна матриця для множини облич не є діагональною.

Для оптимального подання облич необхідно знайти такий простір, в якому коваріаційна матриця для навчальної множини облич буде діагональною. Ця умова буде виконуватися, якщо за базисні вектори простору взяти власні вектори. Базисні вектори простору облич називають основними компонентами. Введемо матрицю Y , що містить основні компоненти, k -ий стовбчик матриці - подання k -го зображення обличчя в просторі облич. Її коваріаційна матриця $\Sigma_Y = Y \cdot Y^T$ буде діагональною.

Для отримання основних компонентів використовуються лінійні перетворення. Нехай P -

матриця перетворення, тоді: $Y=P^T \cdot X$, $X=P \cdot Y$. В [3] запропоновано за матрицю P взяти власні вектори матриці Σ_X : $\Sigma_Y \cdot P = P \cdot \Lambda$, де Λ - діагональна матриця, що містить власні значення матриці Σ_X . Тоді $\Sigma_X = P^T \cdot P \cdot \Lambda = \Lambda$, тобто, Σ_Y - діагональна матриця, що містить власні вектори Σ_X , чого і вимагалось досягнути. Для знаходження власних векторів використовується метод обертань [4].

Величина власного значення визначає важливість або вагу відповідного йому власного вектору для подання простору. Власні значення можна відсортувати за убубанням, та відкинути вектори, що відповідають найменшим власним значенням, не вносячи цим істотної помилки в подальше розпізнавання [3]. При цьому вимірність простору ще зменшиться в декілька раз.

Кожний із знайдених власних векторів визначає точку в просторі зображень; тобто являє собою деяке зображення. Більше того, отримані зображення схожі на обличчя. Тому їх аналогічно до векторів називають "власними обличчями" (ВО), підкреслюючи цим їхній зв'язок зі зображеннями облич. Приклад зображень перших ВО, упорядкованих за убубанням відповідних їм власних значень, наведений на малюнку 1.



Мал. 1. Приклад перших ВО

Зазначимо, що перше ВО є середнім всіх облич, поданих в навчальній множині; друге відповідає напрямку найбільших змін; третє - напрямку наступних по величині змін; і т. д. Зображення будь-якого обличчя може бути уявлене в вигляді лінійної комбінації ВО. На малюнку 2 показаний приклад розкладу одного з облич (тільки перші чотири члена розкладу).



Мал. 2. Приклад розкладу зображення обличчя по базису ВО

Після побудови базису ВО, з'являється можливість подання довільного зображення в знайденому базисі. Причому коефіцієнт розкладу при першому власному векторі може служити детектором наявності людського обличчя на зображенні. Експериментально встановлено, що

коефіцієнт розкладу при першому векторі для зображень, що містять обличчя, складає в середньому 0.96; а для інших зображень - 0.83. Проведені дослідження показали можливість успішного застосування цього методу. Відносна кількість зображень, що були невірно ідентифіковані (обличчя / не обличчя), не перевищувала 2%.

Коефіцієнти розкладу являють собою вектор, вимірність якого дорівнює кількості ВО і звичайно складає декілька десятків. Маючи коефіцієнти розкладу зображення довільного обличчя і знаючи набір ВО, по якому цей розклад виконувався, можна відновити вхідне зображення. На малюнку 3 наведені вхідне і відновлені зображення одного і того ж обличчя при різній кількості коефіцієнтів розкладу. Точність відновлення тим більша, чим більше ВО використовується при розкладі. Тим не менше вже при 50 членах розкладу відновлене зображення мало відрізняється від оригіналу. Таким чином, замість зображень облич можна розглядати вектори в n -мерном просторі.



Мал. 3. Відновлення зображення обличчя людини по різній кількості ВО: а) вхідне зображення; б) - е) відновлене по 100, 50, 25 і 10 ВО відповідно

Для відновлення зображень на мал.3 ВО обчислювалися по всім 212 зображенням (приблизно по 20 зображень облич кожної з 11 людей) і їх було отримано також 212. Після цього отримані ВО ранжувалися в порядку убубання відповідних їм власних значень, і при відновленні використовувалося тільки 100, 50, 25 або 10 перших з них.

Слід підкреслити, що для успішного застосування цього методу зображення облич повинні бути справді "схожі" друг на друга, тобто обличчя повинні мати однаковий масштаб, ракурс і розташуватися в одному і тому ж місці на зображенні. Бажано, щоб обличчя займало все зображення. Це дозволить виключити вплив фону.

2. РОЗПІЗНАВАННЯ ЛЮДСЬКИХ ОБЛИЧ

Подання зображень людських облич як векторів в просторі ВО дає можливість вирішувати задачу розпізнавання облич. Для цього спочатку за декількома десятками зображень довільних облич

формується ВО. Після цього декілька зображень кожного з "відомих" людей розкладаються по цим ВО і вектори коефіцієнтів розкладу зберігаються в базі даних як еталони для цих людей. Зазначимо, що зображення, що розкладаються, могли і не використовуватися при формуванні ВО. База даних еталонних векторів в будь-який момент може бути доповнена новими обличчями. Набір ВО при цьому залишається незмінним. Кожне нове введені зображення обличчя людини також розкладається по тим же самим ВО, що й еталонні зображення.

Еталонні вектори, що отримані розкладом по ВО зображень обличчя кожної людини, утворюють класи. Тепер можна ставити і вирішувати задачу приналежності вектору, отриманого розкладом по ВО зображення довільного обличчя, одному з класів.

2.1. БАЗИ ДАНИХ

В експериментах по розпізнаванню обличчя використовувались зображення з декількох баз даних. Оскільки розміри обличчя на зображеннях різних баз даних істотно розрізняються, експерименти з зображеннями кожної бази даних проводилися незалежно від інших. У базах були представлені обличчя людей різної статі та расової приналежності, з вусами, бородами, окулярами та без них. Варіювався також вираз обличчя. Нижче наведені більш детальні параметри кожної з баз даних.

Манчестерська база даних: 214 зображень розміром 64x64 (20 зображень обличчя кожного з 9 чоловіків і 3 жінок). Мають місце слабкі повороти голови, трохи розрізняється масштаб, на окремих кадрах наявні окуляри.

База даних дослідної лабораторії Оліветті: 400 зображень розміром 92x112 (по 10 зображень обличчя кожного з 35 чоловіків і 5 жінок). На зображеннях однієї й тієї ж людини розрізняються вираз обличчя та ракурс. Масштаби практично однакові.

База даних Йельського університету: 165 зображень розміром 320x243 (по 11 зображень обличчя кожного з 14 чоловіків та 1 жінки). На зображеннях однієї й тієї ж людини розрізняються вираз обличчя, умови освітлення і наявність окулярів. Масштаб та ракурс практично однакові.

База даних університету Берну: 150 зображень розміром 512x342 (по 10 зображень обличчя кожного з 15 чоловіків). На зображеннях однієї й тієї ж людини розрізняються вираз обличчя і ракурс. Масштаб практично однаковий.

База даних MIT: 450 зображень розміром 128x120 (по 27 зображень обличчя кожного з 16 чоловіків). На зображеннях однієї й тієї ж людини

розрізняються вираз обличчя, ракурс, масштаб та умови освітлення. Тільки серед зображень цієї бази є бічні нахили голови.

2.2. РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТІВ

Перед виконанням розпізнавання зображення всіх баз даних були заздалегідь оброблені: фон був обрізаний і залишено тільки обличчя. Для порівняння манчестерська база даних, на зображеннях якої фон слабкий, була оброблена також і без попереднього виділення обличчя.

Експерименти проводилися слідуєчим чином. З зображень бази даних формувалася навчальна вибірка, розмір якої варіювався від всієї бази до одного зображення на людину. По зображенням навчальної вибірки обчислювалися ВО. Після цього зображення кожного з наведених в базі людей ділилися на дві рівні частини. По першій половині обчислювалися еталонні вектори коефіцієнтів розкладу, а зображення другої розглядалися як невідомі і підлягали розпізнаванню. Після цього половини мінялися місцями. Еталонні вектори коефіцієнтів розкладу, стосовні до однієї людини, осереднювалися для отримання середнього вектору.

Кожне нове зображення обличчя невідомої людини розкладалося по тим же самим ВО, що й еталонні зображення. Вважалася, що обличчя на введеному зображенні належить тій людині, чий середній еталонний вектор виявлявся найближчим.

Частки невірно розпізнаних обличчя при різних розмірах навчальної вибірки і кількості ВО наведені в таблиці 1.

Результати експериментів показують, що даний метод розпізнавання людських обличчя надто чутливий до фону на зображенні. В першій серії експериментів зображення манчестерської бази даних були оброблені разом з фоном (перший рядок в табл.1). Частка помилок при розпізнаванні в цьому випадку була дуже висока і досягала половини. Це, звичайно, менше, ніж імовірність помилок при випадковому виборі, але значно більше, ніж можна допустити в реально працюючій системі. Усунення фону дозволило в 4-5 раз знизити частку помилок (третій рядок в табл.1).

Ще один шлях зниження помилок - використання більш досконалої процедури розпізнавання. Так пошук найближчого з всіх еталонних векторів (а не середнього) дозволяє зменшити імовірність помилки в півтора рази (другий рядок в табл. 1). Це означає, що кластери, утворені еталонними векторами окремих людей, не мають сферичної форми і погано рознесені в просторі.

Найкращі результати були отримані для зображень бази даних Йельського університету, незважаючи на те, що зображення обличчя однієї й тієї ж людини розрізнялися виразом, умовами освітлення та наявністю окулярів. Отже ці умови не такі важливі як масштаб обличчя і ракурс голови, які на цих зображеннях практичні однакові.

Більш за все помилок було зареєстровано при розпізнаванні обличчя на зображеннях бази даних MIT, серед яких були наявні бічні нахили голови та істотна зміна розміру обличчя. Справді, у випадку

бічного нахилу голови зображення обличчя сильно викривляється і мало схоже на зображення, де голова не нахилена. Зображення, на яких обличчя мають різні розміри також значно розрізняються.

При обробці зображень з баз даних дослідної лабораторії Оливетти й університету Берну, що містять зображення обличчя при поворотах і вертикальних нахилах, отриман невеликий відсоток помилок, отже такі зміни ракурса припустимі для систем розпізнавання, що основані на поданні людських обличчя через власні обличчя.

Табл.1 Результати експериментів

N	База Даних	Кількість зображень в навчальній вибірці	Кількість ВО	Розмір зображення	Еталони/Невідомі		
					Всі/всі	1/2	2/1
1	Manchester*	212	64	64x64	4.7%	48.1%	50.0%
		66	64		12.3%	49.0%	40.7%
		33	32		13.2%	51.0%	46.3%
		11	11		20.8%	56.7%	50.0%
2	Manchester**	212	64	64x64	0%	35.6%	35.2%
3	Manchester	212	32	32x48	3.3%	15.4%	10.2%
		33	32		3.3%	17.3%	9.3%
		11	11		9.4%	26.9%	13.9%
4	ORL	400	32	92x112	4.0%	21.0%	16.0%
		40			5.5%	18.0%	20.0%
5	Yale	135	32	112x128	2.22%	5.0%	12.0%
6	Bern	150	32	96x112	6.7%	16.0%	13.3%
7	MIT	430	32	64x64	29.1%	34.4%	44.7%

* Манчестерська база даних оброблена без виділення облич.

** При розпізнаванні використовувався найближчий еталонний вектор.

ВИСНОВКИ

Таким чином, розклад зображень людських обличчя на ВО може бути використаний при побудові системи розпізнавання людей за їхніми обличчями.

Методи, що основані на ВО дозволяють розпізнавати людей за їхніми обличчями незалежно від вираза обличчя та умов освітлення.

Цей метод має високу чутливість до навколишнього фону, розміру обличчя на зображенні, а також до бічних нахилів голови. При створенні реально працюючої системи, що розпізнає, необхідно передбачити автоматичне визначення положення, розміру і нахилу обличчя на зображенні з тим, щоб компенсувати викривлення, які вносяться, і усувати з розгляду навколишній фон.

Слід також приділити особливу увагу застосуванню ефективних засобів вирішення задачі розпізнавання образів, що здатні істотно підвищити якість розпізнавання.

ЛІТЕРАТУРА

1. Turk M. A., Pentland A. P. Face Recognition Using Eigenfaces.//Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, 1991.-p. 586-591.
2. Pentland A., Moghaddam B., Starner T., Turk M. View-based and Modular Eigenfaces for Face Recognition.//Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, 1994.
3. Romdhani S. Face Recognition Using Principal Component Analysis./The Msc Thesis, 1996.-57p.
4. Березин И.С., Жидков Н.П. Методы вычислений. - М., Наука, 1965.