

УДК 004.93'12:57.087.1(045)

МЕТОД ЗАБЕЗПЕЧЕННЯ ІНВАРІАНТНОСТІ ЗОБРАЖЕННЯ ОБЛИЧЧЯ ЩОДО АФІННИХ СПОТВОРЕНЬ

Д. М. Федоров

Національний авіаційний університет

vorodef@ukr.net

Проаналізовано літературні джерела з розпізнавання зображень облич. Виявлено, що при розпізнаванні зображень враховуються не всі види можливих спотворень, за яких прямі переходять у прямі. Запропоновано враховувати при розпізнаванні облич спотворення, яке полягає у стисненні або розтягненні зображення обличчя і при цьому границі зображення набувають форми паралелограма. Розроблено параметричну формулу для спотвореного таким чином зображення, а також запропоновано правило класифікації, за яким обчислюватиметься міра схожості поданого зображення з еталоном.

Ключові слова: розпізнавання образів, ідентифікація особи, інваріантність зображення, афінні спотворення, інтелектуальні системи, класифікація зображень.

There are analyzed the literature of recognition of face images. There is found that the image recognition is not taken into account all the possible types of distortion, in which straight moving in straight lines. There are considered when faces are detected distortion, which is compressing or stretching the face image and thus limit images take the form of a parallelogram. There is designed the parametric formula for distorted so images, and proposed classification rule, which will be calculated measure of similarity is a representation of the standard.

Keywords: pattern recognition, person identification, image invariance, affine distortion, intelligent systems, classification of images.

Вступ

На сьогодні проблема розпізнавання облич є дуже актуальною. В багатьох інтелектуальних системах, у яких використовується ідентифікація осіб, застосовуються алгоритми розпізнавання облич за фотознімком. Завдання ідентифікації облич та ідентифікація будь-яких об'єктів у цілому складається з двох частин: навчання та розпізнавання. Навчання здійснюється шляхом показу окремих об'єктів із зазначенням їх належності до того чи іншого класу. Зрозуміло, що у випадку облич – обличчя однієї особи може бути тільки одним, але його зображення може бути з незначними спотвореннями, як наприклад невеликий нахил вліво чи вправо або зміщення по осі абсцис та ординат. У результаті навчання розпізнавальна система повинна мати здатність реагувати однаковими реакціями на всі об'єкти одного образу і різними — на всі об'єкти різних образів. Важливо, що в процесі навчання вказуються тільки самі об'єкти та їх належність до образу. За навчанням іде процес розпізнавання нових об'єктів навченою системою. Автоматизація цих процедур і становить проблему навчання розпізнаванню образів. У тому випадку, коли спеціаліст із штучного інтелекту сам розробляє, а потім задає машині правило класифікації, проблема розпізнавання вирішується частково, оскільки основну і головну частину проблеми навчання людина бере на себе. Проблема навчання розпізнаванню

образів цікава як з прикладної, так і з принципової точки зору. З прикладної точки зору розв'язок цієї проблеми важливий насамперед тому, що він відкриває можливість автоматизувати багато процесів, які до цих пір пов'язували лише з діяльністю живого мозку. Принципове значення проблеми тісно пов'язане з питанням, яке все частіше виникає у зв'язку з розвитком ідей кібернетики: що може і що принципово не може робити машина? Зокрема, чи може машина розвинути в собі здатність перейняти в людини вміння виконувати певні дії залежно від ситуацій, що виникають у навколишньому середовищі? Поки що стало зрозуміло тільки те, що якщо людина може спочатку сама усвідомити своє вміння, а потім його описати, тобто вказати, чому виконуються дії у відповідь на кожний стан зовнішнього середовища або як (за яким правилом) об'єднуються окремі об'єкти в образи, то таке вміння без принципових труднощів може бути передано машині. Якщо ж людина володіє умінням, але не може пояснити його, то залишається тільки один шлях передачі його машині — навчання на прикладах. Коло завдань, які можуть вирішуватися за допомогою розпізнавальних систем, надзвичайно широке. До них належать не тільки завдання розпізнавання облич, а й розпізнавання складних процесів і явищ, що виникають при виборі оптимального управління технологічними,

економічними, транспортними або військовими операціями. При вирішенні завдань управління методами розпізнавання образів замість термінів «зображення» або «відображення» застосовують термін «стан».

Стан — це множина відображень вимірюваних поточних (або миттєвих) характеристик об'єкта, що досліджується в слідкуючій системі. Сукупність станів визначає ситуацію. Поняття «ситуація» є аналогом поняття «образ». Але ця аналогія не повна, оскільки не всякий образ можна назвати ситуацією, хоча всяку ситуацію можна назвати образом. Якщо як об'єкт спостереження розглядається деякий об'єкт управління, то поняття «ситуація» об'єднує такі стани цього об'єкта, в яких слід застосовувати одні й ті самі управляючі дії. Якщо об'єктом спостереження є обличчя людини, яка рухається, то ситуація об'єднує всі положення обличчя, при яких воно розташоване паралельно до камери (або має кут нахилу до поверхні об'єктива не більше від наперед заданого значення). Вибір вихідного опису об'єктів є однією з центральних завдань проблеми розпізнавання образів. При вдалому виборі вихідного опису (простору ознак) задача розпізнавання може виявитися тривіальною і, навпаки, невдало вибраний початковий опис може призвести або до дуже складної подальшої обробки інформації, або взагалі до відсутності розв'язку. Наприклад, якщо розв'язується завдання розпізнавання облич, що відрізняються за кутом нахилу в площині XU , а як вихідний опис обрано значення, які відповідають кутам нахилу в площині XZ , то в такому випадку завдання розпізнавання в принципі не може бути розв'язане.

Аналіз останніх досліджень та публікацій

Як показує практика та літературні джерела, більшість зображень облич людей, які потрібно ідентифікувати, подаються на розпізнавання спотвореними. Це може бути або незначний поворот навколо якої-небудь точки зображення, або сильне спотворення зображення з використанням всіх можливих афінних перетворень.

Афінні перетворення — це відображення площини або простору в себе, за яких прямі переходять у прямі. При афінних перетвореннях зображення може масштабуватись, зміщуватись та обертатись навколо якої-небудь точки. Також, при афінних перетвореннях прямокутники можуть перетворюватись у паралелограми, тобто зображення може стискуватись або розтягуватись, але сторони області, яка обмежує зображення, повинні залишатися паралельними. В розглянутих літературних джерелах описані

методи забезпечення інваріантності зображення обличчя щодо повороту, зміщення та масштабування, але нічого не сказано про стиснення або розтягнення зображення, при якому форма його границь є паралелограмом.

У праці [1] описаний алгоритм розпізнавання, який забезпечує інваріантність зображення обличчя щодо деяких афінних спотворень. У ньому спочатку генерується навчальна вибірка на основі одного еталонного об'єкта, а потім створюється каскад класифікаторів для наявної бази об'єктів. Для цього використовується модуль MATLAB для навчання функцій-класифікаторів (*haartraining*). У нього передається багато параметрів, найважливішими з яких є кількість класифікаторів у каскаді і максимально допустима частота помилкових спрацьовувань. Такий алгоритм машинного навчання називається *AdaBoost* і може використовуватися з іншими алгоритмами навчання. *AdaBoost* є адаптивним алгоритмом, оскільки кожен наступний класифікатор будується за об'єктами, які неправильно класифіковані попередніми класифікаторами. Після кожного виклику оновлюється розподіл ваги, який відповідає важливості кожного з об'єктів навчання. На кожній ітерації вага кожного неправильно класифікованого об'єкта зростає і таким чином новий класифікатор «Фокусує свою увагу» на цих об'єктах і краще їх ідентифікує. Після навчання класифікаторів, каскад готовий до роботи з розпізнавання облич. Але можливостями запропонованого в праці [1] алгоритму є розпізнавання облич зі змінним масштабом і облич, повернутих на кут у межах 30 градусів. Обличчя, які стиснені або розтягнуті і область яких має форму паралелограма, алгоритм не розпізнає.

У джерелах [2; 4] для розпізнавання спотвореного зображення використовуються елементарні лінійні перетворення функції зображення обличчя. Тобто, якщо функція зображення обличчя має вигляд $y = f(x)$, то функція спотвореного зображення буде мати вигляд $y_c = \alpha f(\gamma x + \delta) + \beta$, де α — коефіцієнт розтягнення по осі абсцис; γ — коефіцієнт розтягнення по осі ординат; δ — зміщення по осі абсцис; β — зміщення по осі ординат.

У джерелі [2] коефіцієнти α , γ , δ , β підбираються таким чином, щоб функція класифікації $C(y, y_c)$ прямувала до нуля. Це все дозволяє розпізнати обличчя, спотворення якого полягають у розтягненні по осі абсцис або розтягненні по осі ординат, або зміщенні по осі абсцис, або зміщенні по осі ординат. У праці [2] нічого не сказано про поворот зображення та про

стиснення його, при якому прямокутна область зображення стає паралелограмом.

У працях [3; 5] автори також торкаються проблеми спотворення зображення обличчя, яке подається на розпізнавання. У джерелах [3; 5] враховуються спотворення, які полягають у зміщенні та повороті зображення обличчя. Поворот виконується за допомогою матриці повороту

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

де (x, y) — координати вектора до повороту;

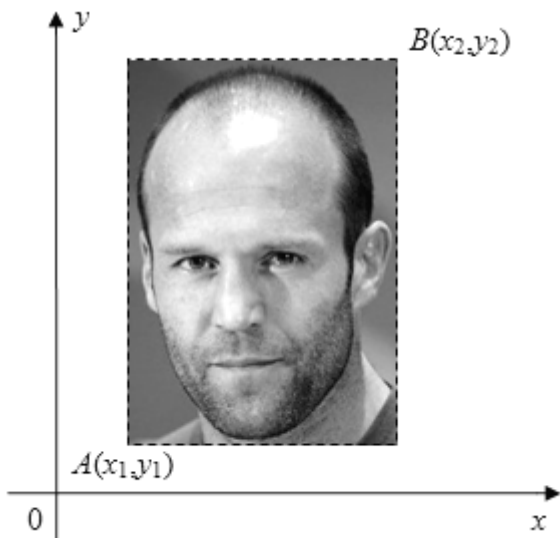
(x', y') — координати вектора після повороту;

θ — кут, на який здійснюється поворот.

Але в роботі більше нічого не сказано про можливі афінні спотворення зображення обличчя, в тому числі і про спотворення типу «паралелограм».

Постановка задачі

Як видно з аналізу останніх досліджень та публікацій, не всі афінні спотворення враховуються при розпізнаванні зображень. Метод нечутливості до спотворення, при якому границі зображення мають форму не прямокутника, а паралелограма, не описаний у розглянутих літературних джерелах. Зображення при такому спотворенні стискається або розтягується по обох координатах, що може негативно впливати на розпізнавання. Тому завданням даної роботи є розроблення параметричної формули для спотвореного таким чином зображення, а також правила класифікації, за яким буде обчислюватись міра схожості поданого зображення з еталоном. Математично це формулюється таким чином: нехай дано оригінальне і спотворене зображення обличчя (рис. 1, а, б).



а

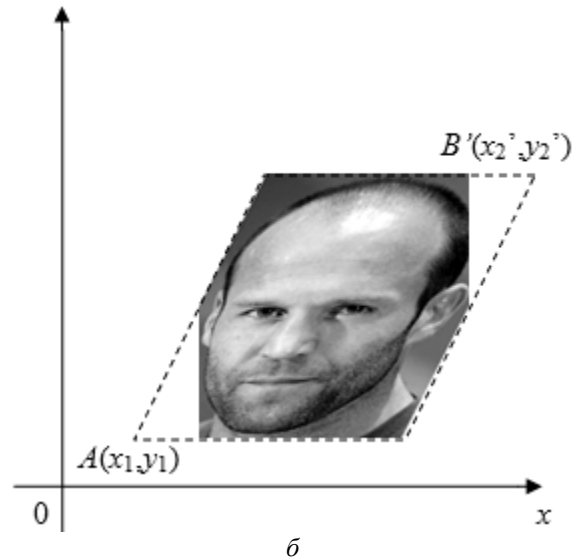


Рис. 1. Зображення обличчя: а — оригінальне; б — спотворене

Якщо функцію оригінального зображення позначити як

$$F(x, y) = z, x_1 \leq x \leq x_2, y_1 \leq y \leq y_2,$$

де z — колір кожного пікселя оригінального зображення, то потрібно знайти параметричну функцію

$$F'(x', y', x_2', y_2') = z, x_1 \leq x \leq x_2', y_1 \leq y \leq y_2',$$

де x_2', y_2' — координати правого верхнього кута області спотвореного зображення.

Також потрібно запропонувати метод пошуку параметрів x_2', y_2' таких, що

$$\sum |G(x, y) - F'(x', y', x_2', y_2')| \rightarrow \min,$$

де $G(x, y)$ — функція поданого зображення.

Розв'язання задачі

На рис. 2 показано прямокутну область оригінального зображення обличчя та область спотвореного зображення у вигляді паралелограма.

З рис. 2 випливає, що кут нахилу паралелограма α можна визначити зі скалярного добутку векторів \overline{DC} і \overline{DC}' , тобто:

$$\begin{aligned} (x_2 - x_2')(x_2' - x_2) + (y_2 - y_2')(y_2' - y_1) &= \\ = \sqrt{(x_2 - x_2')^2 + (y_2 - y_1)^2} \times &\Leftrightarrow \\ \times \sqrt{(x_2' - x_2)^2 + (y_2' - y_1)^2} \cos \alpha & \\ \Leftrightarrow (y_2 - y_1)(y_2' - y_1) = |y_2 - y_1| \times &\Leftrightarrow \\ \times \sqrt{(x_2' - x_2)^2 + (y_2' - y_1)^2} \cos \alpha & \end{aligned}$$

$$\Leftrightarrow \cos \alpha = \frac{(y_2 - y_1)(y_2' - y_1)}{|y_2 - y_1| \sqrt{(x_2' - x_2)^2 + (y_2' - y_1)^2}}. \quad (1)$$

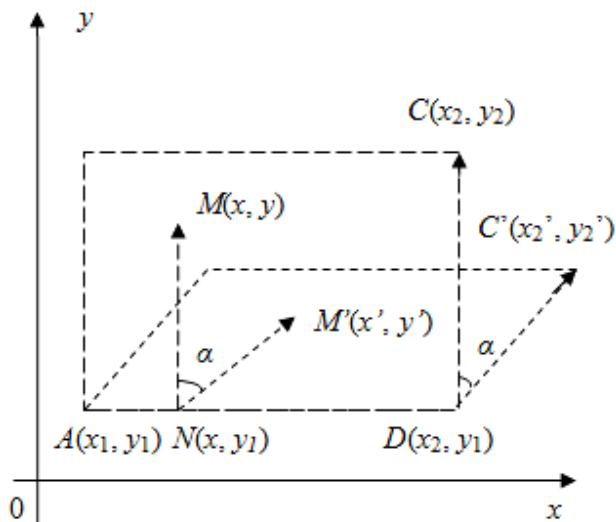


Рис. 2. Облaсті оригiнального та спотвореного зображень

З iншого боку, вектори \overline{NM} i $\overline{N'M'}$ також утворюють мiж собою кут α i можуть бути описанi спiввiдношенням скалярного добутку:

$$\begin{aligned} (x-x)(x'-x) + (y-y_1)(y'-y_1) &= \\ &= \sqrt{(x-x)^2 + (y-y_1)^2} \times \\ &\quad \times \sqrt{(x'-x)^2 + (y'-y_1)^2} \cos \alpha \\ \Leftrightarrow (y-y_1)(y'-y_1) &= |y-y_1| \times \\ &\quad \times \sqrt{(x'-x)^2 + (y'-y_1)^2} \cos \alpha. \end{aligned} \quad (2)$$

Крiм того, оскiльки трикутники MNM' i CDC' подiбнi мiж собою, то мiж ними має мiсце таке спiввiдношення:

$$\frac{|y-y_1|}{\sqrt{(x'-x)^2 + (y'-y_1)^2}} = \frac{|y_2-y_1|}{\sqrt{(x_2'-x_2)^2 + (y_2'-y_1)^2}}. \quad (3)$$

Пiдставимо вираз (1) у вираз (2). Отримаємо:

$$\begin{aligned} (y-y_1)(y'-y_1) &= |y-y_1| \sqrt{(x'-x)^2 + (y'-y_1)^2} \times \\ &\quad \times \frac{(y_2-y_1)(y_2'-y_1)}{|y_2-y_1| \sqrt{(x_2'-x_2)^2 + (y_2'-y_1)^2}}. \end{aligned} \quad (4)$$

З формули (3) випливає, що

$$\begin{aligned} |y-y_1| \sqrt{(x_2'-x_2)^2 + (y_2'-y_1)^2} &= \\ &= |y_2-y_1| \sqrt{(x'-x)^2 + (y'-y_1)^2} \end{aligned}$$

або

$$\frac{\sqrt{(x'-x)^2 + (y'-y_1)^2}}{\sqrt{(x_2'-x_2)^2 + (y_2'-y_1)^2}} = \frac{|y-y_1|}{|y_2-y_1|}. \quad (5)$$

Пiдставляючи формулу (5) у формулу (4) дiстаємо:

$$\begin{aligned} (y-y_1)(y'-y_1) &= \frac{(y-y_1)^2}{(y_2-y_1)^2} (y_2-y_1)(y_2'-y_1) \Rightarrow \\ \Rightarrow y'-y_1 &= \frac{y-y_1}{y_2-y_1} (y_2'-y_1) \Rightarrow \\ \Rightarrow y' &= y_1 + \frac{y-y_1}{y_2-y_1} (y_2'-y_1). \end{aligned}$$

Для обчислення x' формула буде простiшою. З рис. 2 видно, що x' дорiвнює сумi координати x точки N i довжини проекцiї вектора \overline{NM} на вiдрiзок AD , тобто $x' = y' \operatorname{tg} \alpha + x$ або

$$\begin{aligned} x' &= y' \sqrt{\frac{1}{\cos^2 \alpha} - 1} + x \Rightarrow x' = y' \frac{x_2'-x_2}{y_2'-y_1} + x \Rightarrow \\ \Rightarrow x' &= \left(y_1 + \frac{y-y_1}{y_2-y_1} (y_2'-y_1) \right) \frac{x_2'-x_2}{y_2'-y_1} + x \Rightarrow \\ \Rightarrow x' &= y_1 \frac{x_2'-x_2}{y_2'-y_1} + \frac{(y-y_1)(x_2'-x_2)}{y_2-y_1} + x. \end{aligned}$$

Знаючи формули для обчислення координат x' та y' вiдносно координат x та y , функцiя спотвореного зображення буде мати вигляд

$$\begin{aligned} F'(x', y', x_2', y_2') &= F \left(y_1 \frac{x_2'-x_2}{y_2'-y_1} + \right. \\ &\quad \left. + \frac{(y-y_1)(x_2'-x_2)}{y_2-y_1} + x, y_1 + \frac{y-y_1}{y_2-y_1} (y_2'-y_1) \right) = z. \end{aligned}$$

де $x_1 \leq x \leq x_2'$, $y_1 \leq y \leq y_2'$.

Якщо функцiю поданого на розпiзнавання зображення позначити як $G(x, y)$, то рiзниця мiж ним та заданою параметрично функцiєю еталонного зображення буде такою

$$\begin{aligned} d &= \sum_{i=1}^n \left| G(x_i, y_i) - F \left(y_1 \frac{x_2'-x_2}{y_2'-y_1} + \right. \right. \\ &\quad \left. \left. + \frac{(y_i-y_1)(x_2'-x_2)}{y_2-y_1} + x_i, y_1 + \frac{y_i-y_1}{y_2-y_1} (y_2'-y_1) \right) \right| \rightarrow \min, \end{aligned} \quad (6)$$

де n — кiлькiсть точок зображення.

З формули (6) видно, що якщо оригінальне зображення збігається зі спотвореним, або $x_2 = x_2'$ і $y_2 = y_2'$, то вона набуває вигляду

$$d = \sum_{i=1}^n |G(x_i, y_i) - F(x_i, y_i)| \rightarrow \min.$$

Тобто порівнюється еталонне зображення обличчя, в якому не враховуються спотворення із спотвореним зображенням, яке подано на розпізнавання. Зрозуміло, що чим менше різниця між зображеннями, тим більше вони схожі між собою.

Отже, щоб обчислити найменше значення вищенаведеного виразу, потрібно підібрати відповідні параметри x_2' та y_2' , при яких $d \rightarrow \min$.

Маючи параметрично задану функцію $F'(x, y, x_2', y_2')$ для кожного еталонного обличчя особи, можна обчислити значення d для кожного з них.

Подане на розпізнавання обличчя буде належати тій особі, для якої значення d з набору таких обчислених значень буде найменшим.

Висновки

Отже, як показав аналіз літературних джерел з розпізнавання зображень, при розпізнаванні обличчя враховуються не всі види можливих спотворень, які можна назвати афінними.

В розглянутій літературі не описане спотворення, при якому границі зображення мають форму не прямокутника, а паралелограма і при цьому зображення стиснуте або розтягнуте. Тому, була розроблена параметрична формула, в якій варіювання додаткових параметрів дасть змогу охопити множину спотворених зображень обличчя, а також запропоновано правило класифікації, за яким буде обчислюватись міра схожості поданого зображення з еталоном.

ЛІТЕРАТУРА

1. *Синица А. А.* Обнаружение и локализация лиц в системе видеонаблюдения / А. А. Синица, Л. В. Калацкая. — Минск, 2013. — Электроника инфо, № 2. — С. 24.

2. *М. П. Волченков* Об автоматическом распознавании лиц / М. П. Волченков, И. Ю. Самоненко. — Интеллектуальные системы, № 9. — С. 135.

3. *Буй Тхи Тху Чанг.* Распознавание лиц на основе применения метода Виолы-Джонса, вейвлетс-преобразования и метода главных компонент / Чанг Буй Тхи Тху, Хоанг Фан Нгок, В. Г. Спицын // Известия Томского политехнического университета, № 5, 2012.

4. *Долгов С. В.* Современные системы распознавания человека по изображению лица / С. В. Долгов. — Труды научной сессии МИФИ-2003. — М., 2003.

5. *Пентланд А.* Распознавание лиц для интеллектуальных сред / А. Пентланд, Т. Чаудхари. — Открытые системы. — 2000, № 3.

REFERENCES

1. *Sinitsa A. A.* Detection and localization of persons in the video / A. A. Sinitsa, L. V. Kalatskaya. — Minsk, 2013. — Electronics info, № 2. — With. 24.

2. *Volchenkov M. P.* About automatic face recognition / M. P. Volchenkov, I. Y. Samonenko. — Intelligent Systems, № 9. — With. 135.

3. *Chang Bui Thi Thu.* Face recognition by applying the Viola-Jones veyvlets transform and principal component analysis / Bui Thi Thu Chang, Phan Ngoc Hoang, V. G. Spitzin // Tomsk Polytechnic University, number 5, 2012.

4. *Dolgov C. B.* Modern recognition of human facial image / C. B. Dolgov. — Proceedings of the Scientific Session of the MiFi 2003. — Moscow, 2003.

5. *Pentland A.* Face recognition for smart environments / A. Pentland, T. Chaudhary. — Open Systems. — 2000, № 3.

Стаття надійшла до редакції 27.06.2013