

УДК 004.93, 519.2

MSC 62P30

RESTORATION OF AFFINE TRANSFORMED IMAGES

VIACHESLAV V. DUDAR

Faculty of Cybernetics, Taras Shevchenko National University of Kiev, Kiev, Ukraine,
E-mail: slavko123@ukr.net.

ВІДНОВЛЕННЯ АФІННО-ТРАСФОРМОВАНИХ ЗОБРАЖЕНЬ

ВЯЧЕСЛАВ В. ДУДАР

Факультет кібернетики, Київський національний університет імені Тараса Шевченка,
Київ, Україна, E-mail: slavko123@ukr.net.

ABSTRACT. The article contains the description of the developed framework for working with affine transformed images. The first part includes algorithms for estimation the coefficients of affine transformation between images (works both for black-and-white and gray scale images), and the algorithm for estimation the polynomial transformation (works for grey-scale images). The second part includes the tool for making any image recognition algorithm affine-invariant and the approach for obtaining a broad class of affine invariant measures of images. The third part contains the robust modification of the described algorithms, that makes them stable to outliers. Results of implementation and testing are also listed. Developed algorithms could be effectively used in pattern recognition and machine learning.
KEYWORDS: image affine transformation, affine invariant measures of images, recognition of transformed images, robust affine invariant features, comparison of images with outliers.

РЕЗЮМЕ. Стаття містить опис розробленого фреймворку для роботи з афінно-трансформованими зображеннями. У роботі запропоновано алгоритми для знаходження афінного перетворення між зображеннями (що працюють як з чорно-білими, так і сірими зображеннями), та підхід для отримання широкого класу афінно-інваріантних ознак зображень. Також розглянуто робастну модифікацію алгоритмів. У роботі наведено результати імплементації та тестування. Розроблені алгоритми можуть бути ефективно використані у розпізнаванні образів та машинному навчанні.

КЛЮЧОВІ СЛОВА: афінні перетворення зображень, афінно-інваріантні ознаки зображення, розпізнавання трансформованих зображень, робастні афінно-інваріантні ознаки, порівняння афінно-трансформованих зображень з викидами.

1. ЗАГАЛЬНИЙ ПІДХІД ТА ОГЛЯД ІСНУЮЧИХ МЕТОДІВ

Нехай F — деяке зображення. Розглянемо задачу розпізнавання об'єкта, зображеного на F (припускаємо, що на F зображено тільки один об'єкт). Припустимо також, що тип зображеного об'єкту не змінюється при поворотах, переносах, розтягненнях тощо — будь-яких невірджених афінних перетвореннях. В такому випадку необхідно, щоб результат розпізнавання був інваріантним до афінних перетворень.

Наразі існує декілька підходів для розпізнавання трансформованих зображень. Опишемо основні з них.

1) У навчальну вибірку додатково включають також трансформовані навчальні елементи відповідно до необхідних інваріантів (повороти, переноси тощо) [1, с. 261]. Такий підхід неефективний, оскільки вимагає значної кількості додаткових навчальних елементів.

2) До функції похибки додається регуляризуючий доданок, що штрафує зміни у виводі моделі за умови трансформації вводу. Такий підхід приводить до алгоритму, відомого як «tangent propagation» [1, с. 263]. Недоліком алгоритму є те що він працює лише для малих збурень зображення (наприклад, поворот на малий кут).

3) Інваріантність до трансформацій вбудовано у попередню обробку зображень, для яких обчислюються статистики, інваріантні до, наприклад, афінних перетворень [1, с. 261].

4) Фінальна опція – вбудувати інваріантні властивості в структуру моделі, що здійснює розпізнавання (наприклад, згорткові нейронні мережі [1, с. 267] чи вибір специфічної функції ядра в методі релевантних векторів [1, с. 345]).

У даній статті розглянуто два загальних підходи для досягнення цієї мети. Перший підхід полягає в наступному. Припустимо, що ми маємо зображення G , для якого відомо клас зображеного об'єкта (наприклад, еталонне зображення відомої літери). Тоді серед всіх афінних перетворень $A(F)$ зображення F необхідно вибрати таке, щоб $A(F)$ та G були максимально подібними, і вже після цього застосовувати будь-які алгоритми порівняння $A(F)$ та G . Такий комбінований алгоритм розпізнавання буде афінно-інваріантним. Детально алгоритм описано в розділі 2.

Другий підхід — застосувати до F таке афінне перетворення A_F , щоб результат $A_F(F)$ був афінно-інваріантним, тобто, для довільного афінного перетворення $B(F)$ виконувалось $A_{B(F)}(B(F)) = A_F(F)$. Алгоритм описано в розділі 3.

2. АЛГОРИТМ МОМЕНТІВ ТА КУТА

Відтепер розглядатимемо зображення F та G (задані як інтенсивності сірого в кожному пікселі) як функції щільності неперервних двовимірних випадкових величин. Для зведення функції інтенсивності, заданої дискретно в пікселях з координатами (x_i, y_j) , до неперервної функції $f(x, y)$, заданої всюди, використаємо кусково-лінійну апроксимацію.

Нехай точка (x, y) лежить у прямокутнику $x_i \leq x < x_{i+1}$, $y_j \leq y < y_{j+1}$, вершини якого лежать у вузлах рівномірної сітки пікселів з кроками $h_x = x_{i+1} - x_i$ та $h_y = y_{j+1} - y_j$. Позначимо значення функції у вузлах сітки так: $f_{11} = f(x_i, y_j)$, $f_{12} = f(x_i, y_{j+1})$, $f_{21} = f(x_{i+1}, y_j)$, $f_{22} = f(x_{i+1}, y_{j+1})$. Значення функції в точці (x, y) визначимо таким чином:

$$f(x, y) = (f_2 - f_1) \frac{x - x_i}{h_x} + f_1,$$

де $f_1 = (f_{12} - f_{11}) \frac{y - y_j}{h_y} + f_{11}$, $f_2 = (f_{22} - f_{21}) \frac{y - y_j}{h_y} + f_{21}$. Останній крок — нормалізація отриманого двовимірного розподілу.

Позначимо компоненти двовимірних випадкових векторів F та G як (X_F, Y_F) та (X_G, Y_G) відповідно.

Розглянемо довільне афінне перетворення випадкового вектора F :

$$Z = MF + m,$$

де

$$M = \begin{pmatrix} a & b \\ d & e \end{pmatrix}, \quad m = \begin{pmatrix} c \\ f \end{pmatrix}.$$

Покомпонентний запис:

$$\begin{pmatrix} X_Z \\ Y_Z \end{pmatrix} = \begin{pmatrix} a & b \\ d & e \end{pmatrix} \begin{pmatrix} X_F \\ Y_F \end{pmatrix} + \begin{pmatrix} c \\ f \end{pmatrix}.$$

Необхідно апроксимувати коефіцієнти афінного перетворення таким чином, щоб розподіли випадкових векторів Z та G були максимально подібними.

Прирівняємо математичні сподівання Z та G :

$$E(X_Z) = E(aX_F + bY_F + c) = aE(X_F) + bE(Y_F) + c = E(X_G),$$

$$E(Y_Z) = E(dX_F + eY_F + f) = dE(X_F) + eE(Y_F) + f = E(Y_G).$$

Отримаємо систему з двох лінійних рівнянь:

$$\begin{cases} a\mu_{X_F} + b\mu_{Y_F} + c = \mu_{X_G}, \\ d\mu_{X_F} + e\mu_{Y_F} + f = \mu_{Y_G}. \end{cases}$$

Позначимо $\bar{Z} = Z - E(Z)$ та запишемо вираз для рівності коваріаційних матриць:

$$\Sigma_Z = E(\bar{Z}\bar{Z}^T) = E(M\bar{F}\bar{F}^T M^T) = ME(\bar{F}\bar{F}^T)M^T = \Sigma_G.$$

Отримаємо таку систему рівнянь:

$$\begin{pmatrix} a & b \\ d & e \end{pmatrix} \begin{pmatrix} \sigma_{X_F}^2 & \rho_F \\ \rho_F & \sigma_{Y_F}^2 \end{pmatrix} \begin{pmatrix} a & d \\ b & e \end{pmatrix} = \begin{pmatrix} \sigma_{X_G}^2 & \rho_G \\ \rho_G & \sigma_{Y_G}^2 \end{pmatrix}.$$

Дану систему рівнянь можна спростити, якщо звести зображення F та G до таких, що мають одиничну коваріаційну матрицю. Це можна зробити за допомогою лінійного перетворення. Припустимо, що зображення V має коваріаційну матрицю

$$\begin{pmatrix} \sigma_X^2 & \rho \\ \rho & \sigma_Y^2 \end{pmatrix}.$$

Розглянемо зображення W , задане як $W = A_V V$, де

$$A_V = \begin{pmatrix} \frac{\sigma_Y}{\sqrt{\sigma_X^2 \sigma_Y^2 - \rho^2}} & -\frac{\rho}{\sigma_Y \sqrt{\sigma_X^2 \sigma_Y^2 - \rho^2}} \\ 0 & \frac{1}{\sigma_Y} \end{pmatrix}.$$

Тоді зображення W має одиничну коваріаційну матрицю.

Отже, на першому кроці алгоритму зображення F та G зводяться до $F_1 = A_F F$ та $G_1 = A_G G$ з одиничними коваріаційними матрицями. Тепер потрібно знайти афінне перетворення, яке зводить зображення F_1 до G_1 . Система рівнянь для коваріаційних матриць переписеться так:

$$\begin{cases} a^2 + b^2 = 1, \\ ad + be = 0, \\ d^2 + e^2 = 1. \end{cases}$$

Є дві множини розв'язків цієї системи:

$$\begin{pmatrix} a & b \\ d & e \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}, \quad \begin{pmatrix} a & b \\ d & e \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ -\sin \varphi & -\cos \varphi \end{pmatrix},$$

де $\varphi \in [0, 2\pi]$. Перший розв'язок відповідає повороту (на кут φ проти часової стрілки), другий розв'язок — композиції повороту та симетричного відображення відносно осі O_x .

Для апроксимації кута повороту φ розглянемо значення середнього кута зображень F_1 та G_1 відносно центра, який визначається так:

$$E(\alpha_V) = \arg \left(E \left(\frac{V - E(V)}{\|V - E(V)\|} \right) \right),$$

де V — деяке зображення, $\arg(r)$ — аргумент вектора r .

Поворот зображення навколо центра на кут $\Delta\alpha$ змінює середній кут зображення на $\Delta\alpha$. Розглянемо першу множину розв'язків (тільки поворот без симетричного відображення). В цьому випадку для апроксимації кута φ потрібно знайти різницю середніх кутів $E(\alpha_{G_1})$ та $E(\alpha_{F_1})$:

$$\varphi = E(\alpha_{G_1}) - E(\alpha_{F_1}).$$

Тоді перетворення між зображеннями F_1 та G_1 таке:

$$R(\varphi) = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}.$$

Матриця перетворення між початковими зображеннями F та G :

$$M = A_G^{-1} R(\varphi) A_F.$$

Вектор зсуву:

$$m = \mu_G - M\mu_F.$$

Отже, афінне перетворення між початковими зображеннями F та G має такий вигляд:

$$A \begin{pmatrix} x \\ y \end{pmatrix} = M \begin{pmatrix} x \\ y \end{pmatrix} + m.$$

Для другого випадку (поворот та симетричне відображення) маємо:

$$-(E(\alpha_{F_1}) + \varphi) = E(\alpha_{G_1}),$$

$$\varphi = -(E(\alpha_{F_1}) + E(\alpha_{G_1})).$$

Матриця повороту та симетричного відображення:

$$R_1(\varphi) = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ -\sin \varphi & -\cos \varphi \end{pmatrix}.$$

Загальне перетворення між F та G :

$$A \begin{pmatrix} x \\ y \end{pmatrix} = M \begin{pmatrix} x \\ y \end{pmatrix} + m,$$

де $M = A_G^{-1} R_1(\varphi) A_F$, $m = \mu_G - M\mu_F$.

Отже, для апроксимації афінного перетворення потрібно спочатку створити два нові зображення F_1 та G_1 та знайти їх середні кути.

Опишемо підхід, що дозволяє не створювати зображень F_1 та G_1 , а знайти їх середні кути, використовуючи початкові зображення F та G . Для цього введемо до розгляду модифікацію середнього кута, яка є еківаріантною до лінійних перетворень зображення.

Середній кут деякого зображення V був заданий як аргумент вектора

$$v(V) = E \left(\frac{V - E(V)}{\|V - E(V)\|} \right) = E \left(\frac{V - E(V)}{\sqrt{(V - E(V))^T (V - E(V))}} \right).$$

Вектор $v(V)$ в загальному випадку не є еківаріантним до лінійних перетворень зображення V : $v(AZ) \neq Av(Z)$.

Розглянемо таку модифікацію вектора $v(Z)$:

$$v(V) = E \left(\frac{V - E(V)}{\sqrt{(V - E(V))^T \Sigma_V^{-1} (V - E(V))}} \right),$$

де Σ_V — коваріаційна матриця зображення V . Такий вектор середнього кута є лінійно еківаріантним. Дійсно,

$$\begin{aligned} v(AV) &= E \left(\frac{AV - AE(V)}{\sqrt{(AV - AE(V))^T (A^{-1})^T \Sigma_V^{-1} A^{-1} (AV - AE(V))}} \right) = \\ &= E \left(\frac{A(V - E(V))}{\sqrt{(V - E(V))^T \Sigma_V^{-1} (V - E(V))}} \right) = \\ &= AE \left(\frac{V - E(V)}{\sqrt{(V - E(V))^T \Sigma_V^{-1} (V - E(V))}} \right) = Av(V). \end{aligned}$$

Покажемо тепер, як визначити тип внутрішнього перетворення (поворот чи композиція повороту та відображення). Припустимо, ми обчислили вектор середнього кута $v(F)$. Наша мета — визначити, чи потрібно виконувати симетричне відображення відносно вектора середнього кута. Позначимо через $\text{Pr}_{v(F)}^\perp(q)$ проекцію вектора q на вісь, перпендикулярну до $v(F)$.

Розглянемо вектор:

$$w(F) = E \left(\frac{\text{Pr}_{v(F)}^\perp(Z)}{\text{Pr}_{v(F)}^\perp(Z)} \right).$$

Вектор $w(F)$ перпендикулярний до $v(F)$, його напрямок (за часовою чи проти часової стрілки відносно $v(F)$) визначається тим, яка половина зображення важча відносно лінії, що проходить через центр мас $E(F)$ та паралельна вектору $v(F)$. Щоб визначити, чи потрібно робити симетричне відображення F відносно даної лінії, потрібно обчислити також $w(G)$ та порівняти орієнтації (знаки z -компонент векторних добутків $v(F) \times w(F)$ та $v(G) \times w(G)$).

Отже, алгоритм зводиться до такої послідовності кроків: обчислити математичне сподівання та коваріаційні матриці зображень F та G , використовуючи їх, обчислити матриці A_F та A_G . За новою формулою обчислити $v(F)$ та $v(G)$, після чого обчислити кути $\arg A_F v(F)$ та $\arg A_G v(G)$. Знайти $w(F)$ та $w(G)$ та визначити орієнтації двох зображень, після чого в залежності від орієнтацій використати формули одного з двох випадків для афінного перетворення.

Зауваження 1. Алгоритм не працює для зображень, у яких середній кут не визначено (вектор середнього кута $v(F) = 0$), або вектор перпендикулярного середнього кута $w(F) = 0$ (детальніше про ці випадки у наступному розділі).

Зауваження 2. Якщо зображення має декілька градацій сірого, то існує простіший алгоритм для знаходження афінного перетворення між зображеннями. Інтервал можливих інтенсивностей сірого (частіше це $[0, 255]$) розбивається на n підінтервалів. Для кожного підінтервалу I_k формується зображення F_k , інтенсивність пікселів якого дорівнює 0, крім тих пікселів, інтенсивність яких потрапила в інтервал I_k (для них інтенсивність дорівнює початковій). Так отримуються шари зображення F . Аналогічно будуються шари зображення G . При афінному перетворенні кожен шар F_k переходить у відповідний шар G_k . Для кожного з шарів знаходиться центр і записується лінійна система рівнянь, що пов'язує центри шарів з невідомими коефіцієнтами афінного перетворення. Система розв'язується методом найменших квадратів з використанням регуляризації. Цей метод поширюється і на випадок поліноміальних перетворень — отримана система рівнянь, що пов'язує центри, також буде лінійною відносно невідомих коефіцієнтів перетворення. Отже, якщо є два зображення F і G , і відомо, що одне зображення є точним поліноміальним перетворенням (і відома степінь) іншого, тоді даним методом можна знайти коефіцієнти цього перетворення.

Зауваження 3. Алгоритм моментів і кута має перевагу в тому, що він працює і для чорно-білих зображень (без відтінків сірого, що фактично репрезентують множини).

3. ГЕНЕРАЦІЯ АФІННО-ІНВАРІАНТНИХ ОЗНАК ЗОБРАЖЕННЯ

Покажемо, як застосувати підхід, описаний в розділі 2, для генерації широкого класу афінно-інваріантних ознак зображень. Описаний алгоритм може бути застосовано для зведення будь-якого алгоритму розпізнавання до класу афінно-інваріантних.

Розглянемо бінарне відношення $B(F, G)$ на класі всіх двовимірних зображень (двовимірних розподілів) визначене так: $B(F, G) = 1$ тоді і тільки тоді, коли існує не вироджене афінне перетворення таке що $G \stackrel{d}{=} AF + b$, тобто, зображення F може бути зведене до зображення G афінним перетворенням. Це відношення є відношенням еквівалентності. Воно розбиває множини зображень на класи еквівалентності. Всередині кожного класу будь-які зображення можуть бути зведені один до одного афінними перетвореннями. Наше мета — з кожного класу еквівалентності вибрати представника та створити алгоритм, що перетворює вхідне зображення на представника класу еквівалентності, використовуючи афінне перетворення.

Розглянемо зображення F , що задовольняють таким властивостям (далі називатимемо їх властивостями репрезентативного зображення, а зображення називатимемо репрезентативними):

- 1) $\mu_F = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$;
- 2) $\Sigma_F = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$;
- 3) $\arg E \left(\frac{F}{\sqrt{F^T F}} \right) = 0$;
- 4) $\text{sign} E \left(\frac{F_Y}{|F_Y|} \right) = 1$.

Отже, це зображення з центром в точці $(0, 0)$, одиничною коваріаційною матрицею, середнім кутом 0 (це означає, що середній вектор кута колінеарний $(1, 0)$) та верхньою частиною зображення, важчою ніж нижня частина зображення (відносно осі O_x).

Розглянемо також класи еквівалентності з такими властивостями (зауважимо, що ці властивості виконуються або не виконуються одночасно для всіх елементів класу), назвемо їх властивостями не виродженого класу:

- 1) Визначено середній кут (це означає що вектор середнього кута зображення ненульовий).
- 2) Визначено перпендикулярний середній кут (отже, вектор $w(F) = E \left(\frac{\text{Pr}_{v(F)}^\perp(Z)}{\text{Pr}_{v(F)}(Z)} \right)$ не є нульовим, де $v(F)$ — вектор середнього кута).

Неважко перевірити (використовуючи частину 2 — алгоритм моментів та кута), що кожен клас еквівалентності, який задовольняє властивості не виродженого класу, містить єдиний елемент, що задовольняє властивості репрезентативного зображення. Щоб показати, що такий елемент існує, достатньо в системі рівнянь (в алгоритмі розділу 2) прирівняти всі значення моментів до описаних у властивостях репрезентативного зображення, та

знайти явний вигляд перетворення, що зводить дане зображення до репрезентативного (нижче це буде зроблено). Щоб показати, що такий елемент єдиний для кожного класу, необхідно припустити, що таких елементів два, записати афінне перетворення між ними, записати систему рівнянь для моментів та показати, що єдиний її розв'язок — одиничне перетворення.

Знайдемо явний вигляд перетворення, що трансформує вхідне зображення у репрезентативне. Припустимо, ми маємо вхідне зображення F (нагадаємо, що всі зображення в даній статті трактуються як двовимірні випадкові вектори) із середнім $\mu_F = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}$, коваріаційною матрицею, заданою як $\Sigma_F = \begin{pmatrix} \sigma_x^2 & \rho \\ \rho & \sigma_y^2 \end{pmatrix}$, та таким вектором середнього кута: $v_F = E \left(\frac{F - \mu_F}{\sqrt{(F - \mu_F)^T \Sigma_F^{-1} (F - \mu_F)}} \right)$. Ще одна необхідна величина — це орієнтація вектора $w(F) = E \left(\frac{\text{Pr}_{v(F)}^\perp(Z)}{\text{Pr}_{v(F)}^\perp(Z)} \right)$ відносно $v(F)$, що задається так:

$$s(F) = -\text{sign}(w_x v_y - w_y v_x).$$

Величину $w(F)$ може бути обраховано в такий спосіб: для кожної точки зображення (x_i, y_j) з інтенсивністю F_{ij} обраховується величина

$$c_{ij} = (x_i - \mu_x) v_y - (y_j - \mu_y) v_x.$$

У випадку $c_{ij} > 0$ ми додаємо F_{ij} до змінної *lowersum*, якщо $c_{ij} < 0$, ми додаємо F_{ij} до змінної *uppersum* (у випадку $c_{ij} = 0$ не виконуємо жодних дій). Після всіх обчислень, якщо

$$\text{uppersum} > \text{lowersum},$$

тоді присвоюємо $s(F) = 1$, якщо

$$\text{uppersum} < \text{lowersum},$$

тоді присвоюємо $s(F) = -1$. У випадку рівності

$$\text{uppersum} = \text{lowersum}$$

алгоритм не може бути застосовано далі, тобто, клас еквівалентності, до якого належить вхідне зображення, не задовольняє властивість 2 невиродженого класу.

Розглянемо матрицю

$$A_F = \begin{pmatrix} \frac{\sigma_y}{\sqrt{\sigma_x^2 \sigma_y^2 - \rho^2}} & -\frac{\rho}{\sigma_y \sqrt{\sigma_x^2 \sigma_y^2 - \rho^2}} \\ 0 & \frac{1}{\sigma_y} \end{pmatrix}.$$

Отримаємо, що зображення $A_F (F - \mu_F)$ має математичне сподівання $(0, 0)$ та одиничну коваріаційну матрицю. Оскільки $|A_F| = \frac{1}{\sqrt{\sigma_x^2 \sigma_y^2 - \rho^2}} > 0$, то таке перетворення не змінює орієнтації зображення. Після цього перетворення новий вектор середнього кута (тут ми використовуємо лінійно еквіваріантний вектор середнього кута, описаний вище) дорівнює $A_F v_F$. Після цього ми обертаємо зображення (зауважимо, що оскільки коваріаційна матриця

одична та математичне сподівання $(0, 0)$, то обертання навколо $(0, 0)$ та симетричні відображення не змінюють математичне сподівання та коваріаційну матрицю) на кут $-\arg(A_F v_F)$ (проти часової стрілки), використовуючи матрицю повороту:

$$R(-\arg(A_F v_F)) = \begin{pmatrix} \frac{(A_F v_F)_x}{\sqrt{(A_F v_F)^T (A_F v_F)}} & \frac{(A_F v_F)_y}{\sqrt{(A_F v_F)^T (A_F v_F)}} \\ -\frac{(A_F v_F)_y}{\sqrt{(A_F v_F)^T (A_F v_F)}} & \frac{(A_F v_F)_x}{\sqrt{(A_F v_F)^T (A_F v_F)}} \end{pmatrix}.$$

Використовуючи рівність $A_F^T A_F = \Sigma^{-1}$, отримуємо спрощений вигляд матриці повороту:

$$R(-\arg(A_F v_F)) = \frac{1}{\sqrt{v_F^T \Sigma^{-1} v_F}} \begin{pmatrix} (A_F v_F)_x & (A_F v_F)_y \\ -(A_F v_F)_y & (A_F v_F)_x \end{pmatrix}.$$

Отже, після повороту отримуємо зображення:

$$Q = R(-\arg(A_F v_F)) A_F (F - \mu_F),$$

яке має центр мас в початку координат, одиничну коваріаційну матрицю та нульовий середній кут.

Останній крок — виконати симетричне відображення зображення навколо осі O_x у випадку $s(F) = -1$. Результуюче перетворення має такий вигляд:

$$Q = \frac{1}{\sqrt{v_F^T \Sigma^{-1} v_F}} \begin{pmatrix} 1 & 0 \\ 0 & s(F) \end{pmatrix} \begin{pmatrix} (A_F v_F)_x & (A_F v_F)_y \\ -(A_F v_F)_y & (A_F v_F)_x \end{pmatrix} A_F (F - \mu_F).$$

Отже, щоб отримати результуюче зображення, не потрібно будувати проміжних зображень, потрібно лише обчислити середнє, коваріаційну матрицю, лінійно еквіваріантний вектор середнього кута та орієнтацію зображення, і перемножити з матриці.

Отже, даний алгоритм перетворює вхідне зображення до форми, яка є афінно-інваріантною до невироджених афінних перетворень вхідного зображення. Таким чином, щоб зробити будь-який алгоритм розпізнавання зображень афінно-інваріантним, необхідно до вхідного зображення застосувати описаний алгоритм, і отримане зображення використати як вхідне для алгоритму.

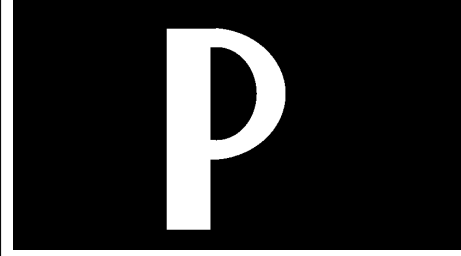


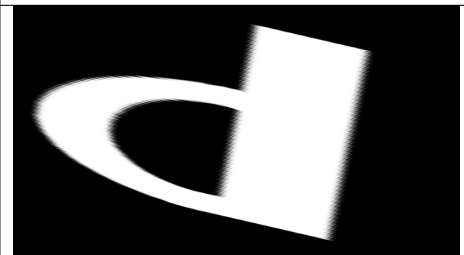


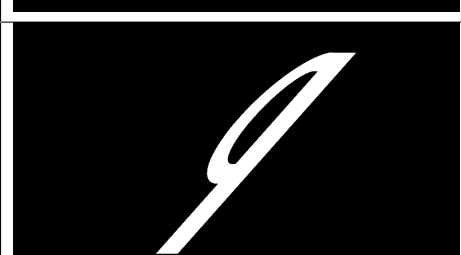

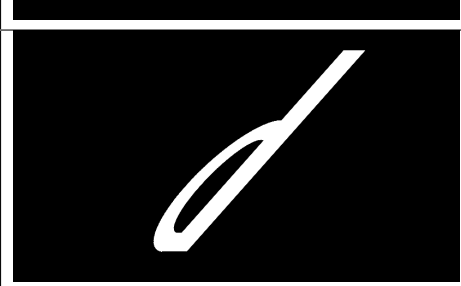

Ще одне застосування алгоритму — генерація широкого класу афінно-інваріантних ознак. Щоб згенерувати афінно-інваріантну ознаку, потрібно застосувати будь-який функціонал Q до результату роботи алгоритму.

Результати тестування алгоритму афінно-інваріантних ознак зображення наведено у таблиці 1.

Описаний вище алгоритм працює лише у випадку, коли клас еквівалентності зображення задовольняє описані вище умови невиродженості — ненульовий середній кут та різна вага половин зображення відносно лінії середнього кута. У випадку, коли середній кут нульовий, можна застосовувати різні модифікації — наприклад, обчислювати середній кут лише для частини розподілу, що знаходиться всередині еліпсу $(Z - m)^T \Sigma^{-1} (Z - m) \leq 1$.

Аналогічні модифікації існують і для випадку рівної ваги половин зображення. Але наведений алгоритм принципово не працює у випадку, коли зображення симетричне відносно певної точки чи лінії.

ТАБЛИЦЯ 1. Результати застосування алгоритму афінно-інваріантних ознак зображення

Вхідне зображення	Вихідне зображення
	
	
	
	
	

4. РОБАСТНИЙ АЛГОРИТМ АФІННО-ІНВАРІАНТНИХ ОЗНАК

Недолік описаних вище алгоритмів — нестійкість до викидів та аномальних спостережень. Причина нестійкості — у використанні математичного сподівання, коваріаційної матриці та середнього кута, які є нестійкими до викидів. Замінімо кожен з цих операторів на стійкий так, щоб загальна схема алгоритму залишилася без змін.

Використаємо багатовимірну медіану замість математичного сподівання. Для алгоритму потрібна медіана, яка є еквіваріантною афінним перетворенням, тобто, для якої виконується рівність $M(AZ + b) = AM(Z) + b$.

Розглянемо таку медіану [2, с. 2] багатовимірної випадкової величини Z :

$$M(Z) = \arg \min_{\mu} E \left(\sqrt{(Z - \mu)^T \Sigma^{-1} (Z - \mu)} \right),$$

де Σ — коваріаційна матриця. Така медіана є афінно-еквіваріантною. Дійсно,

$$\begin{aligned} M(Az + b) &= \arg \min_{\mu} E \left(\sqrt{(AZ + b - \mu)^T (A \Sigma A^T)^{-1} (AZ + b - \mu)} \right) = \\ &= \arg \min_{\mu} E \left(\sqrt{(Z + A^{-1}(b - \mu))^T \Sigma^{-1} (Z + A^{-1}(b - \mu))} \right) = \\ &= A \arg \min_{\mu} E \left(\sqrt{(Z - \mu)^T \Sigma^{-1} (Z - \mu)} \right) + b = AM(Z) + b. \end{aligned}$$

Обчислення такої медіани може бути зведене до обчислення ℓ_1 -медіани. Оскільки Σ^{-1} — симетрична та додатньо визначена, то має місце таке представлення: $\Sigma^{-1} = A^T A$. Тоді

$$\begin{aligned} E \left(\sqrt{(Z - \mu)^T \Sigma^{-1} (Z - \mu)} \right) &= E \left(\sqrt{(Z - \mu)^T A^T A (Z - \mu)} \right) = \\ &= E \left(\sqrt{(AZ - A\mu)^T (AZ - A\mu)} \right). \end{aligned}$$

Позначаючи $Q = AZ$, отримаємо:

$$M(Z) = E \left(\sqrt{(Z - \mu)^T \Sigma^{-1} (Z - \mu)} \right) = A^{-1} M_1(Q),$$

де M_1 — ℓ_1 -медіана. Отже, для обчислення афінно-еквіваріантної медіани необхідно:

- 1) Обчислити матрицю A таку, що $A^T A = \Sigma^{-1}$;
- 2) Перетворити зображення, використовуючи матрицю A : $Q = AZ$;
- 3) Обчислити ℓ_1 -медіану зображення Q ;
- 4) Домножити результат на A^{-1} зліва.

Для обчислення ℓ_1 -медіани використовується ітеративна схема Вейсфельда [3, 4, 6, 5]:

$$m_{k+1} = \sum_{i=1}^N \sum_{j=1}^M \frac{w_{ij}(x_i, y_j)}{\|m_k - (x_i, y_j)\|} \cdot \left(\sum_{i=1}^m \frac{w_{ij}}{\|m_k - (x_i, y_j)\|} \right)^{-1}.$$

Тут (x_i, y_j) — координати пікселів, w_{ij} — інтенсивність пікселя (x_i, y_j) , N , M — розмірність зображення. Щоб уникнути ділення на числа, близькі до нуля, знаменник $\|m_k - x_i\|$ замінюється на $\max(\|m_k - x_i\|, \varepsilon)$. Ітерації припиняємо, якщо $\|m_k - m_{k-1}\| < \varepsilon$. Останній отриманий вектор m_k приймаємо за значення медіани. Надалі медіану $M(F)$ позначатимемо також m_F .

Покажемо тепер, як отримати робастну версію коваріаційної матриці, що змінюється під впливом афінних перетворень випадкової величини в такий самий спосіб, як і звичайна коваріаційна матриця. Це означає, що афінне перетворення випадкової величини $Z \rightarrow AZ + \mu$ змінює матрицю так: $\Sigma_Z \rightarrow A\Sigma_Z A^T$.

Нехай маємо зображення F , для якого знайдено афінно-еквіваріантну медіану $m_F = (m_{F_x}, m_{F_y})$. Побудуємо такий тривимірний випадковий вектор:

$$C_F = \begin{pmatrix} (F_x - m_{F_x})^2 & (F_x - m_{F_x})(F_y - m_{F_y}) \\ (F_x - m_{F_x})(F_y - m_{F_y}) & (F_y - m_{F_y})^2 \end{pmatrix} = \begin{pmatrix} \overline{F_x^2} & \overline{F_x F_y} \\ \overline{F_x F_y} & \overline{F_y^2} \end{pmatrix}$$

і знайдемо 3-вимірну афінно-еквіваріантну медіану цього вектора:

$$M(C_F) = \begin{pmatrix} m_{F_{xx}} \\ m_{F_{xy}} \\ m_{F_{yy}} \end{pmatrix}.$$

Вектор $M(C_F)$ будемо асоціювати з матрицею

$$M_{C_F} = \begin{pmatrix} m_{F_{xx}} & m_{F_{xy}} \\ m_{F_{xy}} & m_{F_{yy}} \end{pmatrix}.$$

Дана матриця є аналогом звичайної коваріаційної матриці. Покажемо, що матриця M_{C_F} змінюється під впливом афінних перетворень таким же чином, як і коваріаційна матриця. Припустимо, що випадковий вектор F змінено на $F_1 = AF + \mu$:

$$\begin{pmatrix} F_{1x} \\ F_{1y} \end{pmatrix} = \begin{pmatrix} aF_x + bF_y + \mu_x \\ dF_x + eF_y + \mu_y \end{pmatrix}.$$

Оскільки медіана є афінно-еквіваріантною, то

$$M(F_1) = AM(F) + \mu,$$

$$\begin{pmatrix} m_{F_{1x}} \\ m_{F_{1y}} \end{pmatrix} = A \begin{pmatrix} m_{F_x} \\ m_{F_y} \end{pmatrix} + \mu = \begin{pmatrix} am_{F_x} + bm_{F_y} + \mu_x \\ dm_{F_x} + em_{F_y} + \mu_y \end{pmatrix}.$$

Тоді отримаємо:

$$\begin{aligned}
 C(F_1) &= \begin{pmatrix} \frac{\overline{F_{1x}}^2}{F_{1x} \overline{F_{1y}}} \\ \frac{\overline{F_{1y}}^2}{F_{1y} \overline{F_{1x}}} \end{pmatrix} = \begin{pmatrix} (F_{1x} - m_{F_{1x}})^2 \\ (F_{1x} - m_{F_{1x}})(F_{1y} - m_{F_{1y}}) \\ (F_{1y} - m_{F_{1y}})^2 \end{pmatrix} = \\
 &= \begin{pmatrix} (a\overline{F_x} + b\overline{F_y})^2 \\ (a\overline{F_x} + b\overline{F_y})(d\overline{F_x} + e\overline{F_y}) \\ (d\overline{F_x} + e\overline{F_y})^2 \end{pmatrix} = \\
 &= \begin{pmatrix} a^2\overline{F_x}^2 + 2ab\overline{F_x}\overline{F_y} + b^2\overline{F_y}^2 \\ ad\overline{F_x}^2 + (ae + bd)\overline{F_x}\overline{F_y} + be\overline{F_y}^2 \\ d^2\overline{F_x}^2 + 2de\overline{F_x}\overline{F_y} + e^2\overline{F_y}^2 \end{pmatrix} = \\
 &= \begin{pmatrix} a^2 & 2ab & b^2 \\ ad & ae + bd & be \\ d^2 & 2de & e^2 \end{pmatrix} \begin{pmatrix} \overline{F_x}^2 \\ \overline{F_x}\overline{F_y} \\ \overline{F_y}^2 \end{pmatrix}.
 \end{aligned}$$

Отже, вектор $C(F_1)$ змінюється лінійно під впливом афінного переворення $AF + \mu$. Оскільки медіана $M(C(F_1))$ є афінно-еквіваріантною, то вона змінюється так:

$$M(C(F_1)) = \begin{pmatrix} a^2 & 2ab & b^2 \\ ad & ae + bd & be \\ d^2 & 2de & e^2 \end{pmatrix} M(C(F)).$$

Тоді матриця $M_{C_{F_1}}$, що асоціюється з вектором $M(C(F_1))$, дорівнює такому виразу:

$$\begin{pmatrix} a & b \\ d & e \end{pmatrix} \begin{pmatrix} m_{F_{xx}} & m_{F_{xy}} \\ m_{F_{xy}} & m_{F_{yy}} \end{pmatrix} \begin{pmatrix} a & d \\ b & e \end{pmatrix} = AM_{C_F}A^T.$$

Отже, матриця M_{C_F} змінюється під впливом афінних перетворень випадкового вектора так, як і звичайна коваріаційна матриця.

Покажемо, як змінити обчислення середнього кута. По-перше, середній кут обчислюється відносно медіани $M(F)$, а не центра мас зображення $E(F)$. По-друге, у формулі середнього кута коваріаційну матрицю Σ замінимо на обчислену робастну коваріаційну матрицю M_{C_F} . Тобто, робастний вектор середнього кута обчислюється за формулою:

$$v_{robust}(F) = E \left(\frac{(F - M(F))}{\sqrt{(F - M(F))^T M_{C_F}^{-1} (F - M(F))}} \right).$$

Ще одна можлива модифікація алгоритму — використання обчисленої робастної коваріаційної матриці M_{C_F} для переобчислення медіани зображення:

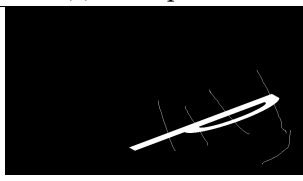

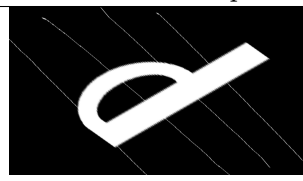
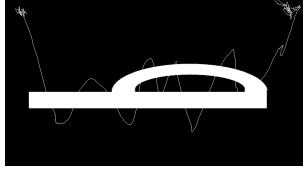
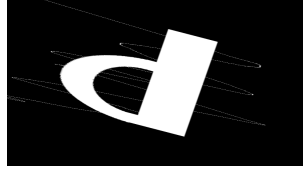

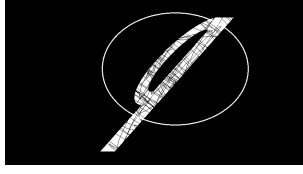
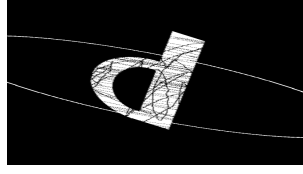
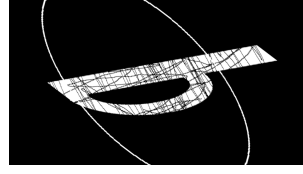



$$M(F)_{recomputed} = \arg \min_{\mu} E \left(\sqrt{(F - \mu)^T M_{C_F}^{-1} (F - \mu)} \right).$$

Використовуючи переобчислену медіану, ми можемо переобчислити робастну коваріаційну матрицю. Під час кожної такої ітерації обчислюються медіани та робастні коваріаційні матриці, які є стійкішими до викидів. Тестування показало, що після двох ітерацій значення медіан та матриць вже практично не змінюються.

Алгоритм афінно-інваріантних ознак задавав у кожному класі еквівалентності одне зображення, яке визначалось набором властивостей (нульове середнє, одинична коваріаційна матриця тощо). Описаний робастний алгоритм також задає такі еталонні зображення (з аналогічними властивостями — нульова медіана, одинична робастна коваріаційна матриця тощо), але тепер малі зміни вхідного зображення (аномальні спостереження, викиди) не спричиняють значних змін у вихідних зображеннях (еталонних зображеннях класів).

У таблиці 2 наведено результати тестування робастного алгоритму афінно-інваріантних ознак зображень.

Таблиця 2. Результати тестування робастного алгоритму афінно-інваріантних ознак зображення

Вхідне зображення	Робастний алгоритм	Звичайний алгоритм
		
		
		
		

ВИСНОВКИ

Розроблений алгоритм афінно інваріантних ознак перетворює зображення до форми, яка не залежить від афінних перетворень вхідного зображення. Це дозволяє, по-перше, генерувати афінно інваріантні ознаки зображень. По-друге, цей підхід можна застосувати для зведення алгоритмів розпізнавання зображень до класу афінно-інваріантних. Розроблена також робастна версія даних алгоритмів, яка є стійкою до викидів та аномальних спостережень на вхідних зображеннях, але вимагає більше обчислень.

ЛІТЕРАТУРА

1. Bishop C. M. Pattern Recognition and Machine Learning. — Springer Science + Business Media, 2006. — 703 p.
2. Hubert Mia, Rousseeuw Peter J., Van Aelst S. High-Breakdown Robust Multivariate Methods. — Statistical Science, 2008. — 29 p., DOI: 10.1214/088342307000000087.
3. Fritz H., Filzmoser P., Croux C. A comparison of algorithms for the multivariate L1-median // <http://www.statistik.tuwien.ac.at/forschung/CS/CS-2010-4complete.pdf>.
4. Kuhn H. W. A note on Fermat's problem // Mathematical Programming. — 1973, Vol. 4. — P. 98–107, DOI: 10.1007/BF01584648.
5. Chandrasekaran R., Tamir A. Open questions concerning Weiszfeld's algorithm for the Fermat-Weber location problem // Mathematical Programming. — 1989, Vol. 44. — P. 293–295, DOI: 10.1007/BF01587094.
6. Aftab K., Hartley R., Trumpf J. Generalized Weiszfeld Algorithms for Lq Optimization // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2015, Vol. 37. — №4. — P. 728–745, DOI: 10.1109/TPAMI.2014.2353625.

Надійшла 10.03.2016