

УДК 519.68

В. М. ГЛАДКИХ, канд. техн. наук,
Одеська національна академія зв'язку ім. О. С. Попова

КОЛІРНИЙ ПРОСТІР ДЛЯ ВИДІЛЕННЯ ТЕКСТУ НА ЗОБРАЖЕННІ ДОКУМЕНТА СУВОРОЇ ЗВІТНОСТІ

Досліджено модель колірному простору перетворення Карунена–Лосва для друкованих бланків документа суворой звітності, заповненого кольоровим рукописним текстом.

Ключові слова: бланк документа; перетворення Карунена–Лосва; сегментація зображення.

Вступ

Залежно від сфери застосування кожне із завдань автоматизації документообігу потребує розробки специфічних методів сегментації (виділення окремих рядків, слів і символів; подання рядків, слів і символів у форматі, необхідному для навчання та класифікації). Особливо це стосується розпізнавання документів суворой звітності, наприклад паспортно-візової інформації, автоматизації обробки поштової кореспонденції тощо.

Особливість упровадження автоматизації оброблення документів суворой звітності полягає в тому, що з метою спрощення відповідних процедур не висуваються жорсткі вимоги стосовно того, чорнилом якого кольору має заповнюватися бланк. У свою чергу, самі бланки можуть бути надруковані або чорною, або будь-якою кольоровою друкарською фарбою. А оскільки бланк документа суворой звітності має, як правило, специфічну структуру, то це практично унеможливило використання відомих методів виділення текстів у документах.

Отже, постає проблема створення специфічних методів виділення рукописного тексту на бланках зазначених документів.

Мета цієї статті — побудова колірному простору та розробка методу сегментації, що дозволяє виділяти на зображенні бланка рукописний текст довільного кольору.

Для сегментації зображень найчастіше використовують простори кольорів RGB, HSI, CIE, Lab, Luv та $I_1I_2I_3$. Проте жодний із них не має таких переваг, що забезпечують високоякісну сегментацію всіх видів кольорових зображень. Тому вибір колірному простору для чіткої сегментації певного класу зображень — завдання, ще й досі є актуальне [1; 2].

Припустимо, що колірний контент власне бланка являє собою градації сірого, а напис кольоровий. При скануванні в кольоровому режимі цифрове зображення бланка може містити спотворення кольорів, проте, як було показано [3], це не має істотного впливу на колірний контент зображення, коли необхідно класифікувати бланк за його контентом у просторі кольорів. Що ж до завдання сегментації цифрового зображення документа суворой звітності, то воно вимагає проведення додаткового дослідження.

Для розробки методу сегментації зображення документа проаналізуємо вплив його колірному контенту на власні значення та власні вектори коваріаційної матриці, розрахованої для компонентів простору RGB.

Елементи коваріаційної матриці S відсканованого зображення документа визначаються за формулами:

$$S_{rr} = \sum_{j=1}^n \sum_{k=1}^m r_{ij}^2 - \bar{r}^2, \quad S_{gg} = \sum_{j=1}^n \sum_{k=1}^m g_{ij}^2 - \bar{g}^2, \quad S_{bb} = \sum_{j=1}^n \sum_{k=1}^m b_{ij}^2 - \bar{b}^2, \quad (1)$$

$$S_{rg} = S_{gr} = \sum_{j=1}^n \sum_{k=1}^m r_{ij} g_{ij} - \bar{r} \bar{g}, \quad S_{rb} = S_{br} = \sum_{j=1}^n \sum_{k=1}^m r_{ij} b_{ij} - \bar{r} \bar{b}, \quad S_{gb} = S_{bg} = \sum_{j=1}^n \sum_{k=1}^m b_{ij} g_{ij} - \bar{g} \bar{b}. \quad (2)$$

Тут m і n — кількість відповідно рядків і стовпців у зображенні; r_{ij} , g_{ij} , b_{ij} — відповідно r -, g -, b -компоненти пікселя в j -му рядку та k -му стовпці; \bar{r} , \bar{g} , \bar{b} — глобальні середні значення RGB-компонентів по всьому зображенню.

Власні значення та вектори — це розв'язки характеристичного рівняння

$$\det(S - \lambda E) = 0,$$

тоді як власні вектори U відповідають власним значенням, які знаходимо, розв'язуючи систему рівнянь

$$(S - \lambda E)U^T = 0.$$

У цих рівняннях E — одинична матриця; $U^T = (u_1, u_2, u_3)$ — матриця-рядок.

Колірний контент відсканованого зображення документа складається з таких кластерів: тло, близьке до білого, — майже 90% пікселів із близькими за значеннями компонентами RGB, $r, g, b \in (240, 255)$; бланк, надрукований друкарськими фарбами, близькими до чорного кольору, — майже 9% пікселів, також із близьким за значеннями компонентами RGB, $r, g, b \in (0, 64)$; кольоровий напис — 1% пікселів зображення. Найчастіше бланк заповнюється чорнилом, близьким до синього кольору.

З огляду на це біле тло — позначимо цей кластер як C_0 , можна подати однаковими середніми значеннями компонентів RGB: $r = g = b = m_0$; друкарський текст бланка — кластер C_1 , подати значеннями $r = g = b = m_1$, а напис C_2 (наприклад, у разі синього кольору) — значенням $r = g = k_0$, $b = k_2$.

Тоді глобальний середній компонент простору RGB визначатиметься зі співвідношень:

$$\begin{aligned}\bar{r} &= m_0 - (m_0 - m_1) \frac{N_1}{N} - (m_0 - k_0) \frac{N_2}{N}; \\ \bar{g} &= m_0 - (m_0 - m_1) \frac{N_1}{N} - (m_0 - k_0) \frac{N_2}{N}; \\ \bar{b} &= m_0 - (m_0 - m_1) \frac{N_1}{N} - (m_0 - k_2) \frac{N_2}{N},\end{aligned}$$

де N — загальна кількість пікселів у зображенні; N_1 і N_2 — кількість пікселів відповідно бланка і напису.

Позначимо $t_1 = N_1/N$ і $t_2 = N_2/N_1$. Тоді з урахуванням властивостей коваріаційної матриці та формул (1) і (2), дістанемо для обчислення її елементів такі вирази:

$$S_{rr} = S_{rg} = S_{gg} = t_1[(1 - t_1 - t_2)(\Delta m_1 + t_2 \Delta k_0)^2 + t_2(\Delta k_0 - t_1(\Delta m_1 + t_2 \Delta k_0))^2 + (\Delta m_1(1 - t_1) - t_1 t_2 \Delta k_0)^2]; \quad (3)$$

$$\begin{aligned}S_{rb} = S_{gb} &= t_1[(1 - t_1 - t_2)(\Delta m_1 + t_2 \Delta k_0)(\Delta m_1 + t_2 \Delta k_2) + \\ &+ t_2(\Delta k_0 - t_1(\Delta m_1 + t_2 \Delta k_0))(\Delta k_0 - t_1(\Delta m_1 + t_2 \Delta k_2)) + \\ &+ (\Delta m_1(1 - t_1) - t_1 t_2 \Delta k_0)(\Delta m_1(1 - t_1) - t_1 t_2 \Delta k_2)];\end{aligned} \quad (4)$$

$$S_{bb} = t_1[(1 - t_1 - t_2)(\Delta m_1 + t_2 \Delta k_2)^2 + t_2(\Delta k_2 - t_1(\Delta m_1 + t_2 \Delta k_2))^2 + (\Delta m_1(1 - t_1) - t_1 t_2 \Delta k_2)^2]. \quad (5)$$

Тут $\Delta m_1 = m_0 - m_1$, $\Delta k_0 = m_0 - k_0$, $\Delta k_2 = m_0 - k_2$.

Розглянемо випадок, коли $t_2 \rightarrow 0$, тобто коли колірний контент зображення подається у градаціях сірого. Тоді всі елементи коваріаційної матриці

$$S_{ij} \rightarrow t_1(1 - t_1)\Delta m_1^2,$$

а сама матриця набирає вигляду

$$S = t_1(1 - t_1)\Delta m_1^2 \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}. \quad (6)$$

При цьому власні значення й власні вектори обчислюватимуться за формулами:

$$\begin{aligned}\lambda_1 &= 3t_1(1 - t_1)\Delta m_1^2, \quad \lambda_2 = \lambda_3 = 0, \\ u_1 &= \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right), \quad u_2 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0 \right), \quad u_3 = \left(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right).\end{aligned}$$

Звідси випливає, що в разі зображення з контентом у градаціях сірого його проекції на вектори u_2 та u_3 дорівнюють нулю, а проекція u_1 — зважена сума компонентів простору RGB.

Розглянемо коваріаційну матрицю, елементи якої обчислюються згідно з (3), (4) і (5) за умови, що членами порядку $O(t_2^3)$ можна знехтувати.

Тоді

$$\begin{aligned}S_{rr} = S_{rg} = S_{gg} &\approx t_1(\Delta m_1 + t_2 \Delta k_0 - t_1(1 + t_2)(\Delta m_1 + t_2 \Delta k_0)); \\ S_{rb} = S_{gb} &\approx t_1(\Delta m_1 + t_2(m_0^2 + m_0(k_0 + k_2) + k_0 k_2) - (1 + t_2)(\Delta m_1 + t_2 \Delta k_2)(\Delta m_1 + t_2 \Delta k_0)); \\ S_{bb} &\approx t_1(\Delta m_1 + t_2 \Delta k_2 - t_1(1 + t_2)(\Delta m_1 + t_2 \Delta k_2)).\end{aligned}$$

Уведемо позначення

$$A = t_2 \frac{\Delta k_0}{\Delta m_1}; \quad B = t_2 \frac{\Delta k_2}{\Delta m_1}; \quad c = t_2 \frac{(m_0^2 - m_0(k_0 + k_2)) + k_0 k_2}{\Delta m_1}.$$

У цьому разі ранг коваріаційної матриці дорівнює двом, а сама матриця така:

$$S = \begin{pmatrix} 1+A & 1+A & 1+c \\ 1+A & 1+A & 1+c \\ 1+c & 1+c & 1+B \end{pmatrix}.$$

Власні значення цієї матриці обчислюються за формулами:

$$\lambda_1 = \frac{3+2A+B - \sqrt{(2A+1-B)^2 + 8(c+1)^2}}{2}; \quad \lambda_2 = \frac{3+2A+B + \sqrt{(2A+1-B)^2 + 8(c+1)^2}}{2}; \quad \lambda_3 = 0,$$

а відповідні їм ненормовані власні вектори визначаються так:

$$\bar{u}_1 = (u_{11}, u_{12}, u_{13}) = \left(1, 1, -\frac{2A-B+1 + \sqrt{(2A-B+1)^2 + 8(c+1)^2}}{2(c+1)} \right);$$

$$\bar{u}_2 = (u_{21}, u_{22}, u_{23}) = \left(1, 1, -\frac{2A-B+1 - \sqrt{(2A-B+1)^2 + 8(c+1)^2}}{2(c+1)} \right);$$

$$\bar{u}_3 = (1, -1, 0).$$

Щоб остаточно зробити вибір компонентів простору Карунена–Лосва для сегментації зображення, необхідно виконати аналіз власних векторів та розподілу кольорів залежно від колірної частоти. Вплив колірної частоти на характеристики простору властивостей головних компонентів можна встановити, проаналізувавши їх при різних значеннях k_0 та k_2 . Якщо $k_0 = 0$, а k_2 змінюється від нуля до 255, то колір напису буде змінюватись від чорного до насиченого синього. У разі, коли $k_2 = 225$, а k_0 змінюється від нуля до 255, то колір напису змінюється від насиченого синього до білого. Для того щоб спростити аналіз, уведемо інтенсивність кольору напису, яка обчислюється за формулою

$$I_b = \frac{2k_0 + k_2}{3}.$$

Залежність характеристик простору Карунена–Лосва і θ від I_b

I_b	$m_0 = 255, m_1 = 0$					$m_0 = 255, m_1 = 100$				
	λ_1	λ_2	u_{13}	u_{23}	θ	λ_1	λ_2	u_{13}	u_{23}	θ
0	3,300	0,000	1,00	-2,00	0,00	3,494	0,000	1,00	-2,00	0,00
10	3,280	0,008	0,99	-2,01	0,19	3,461	0,013	0,99	-2,02	0,30
20	3,261	0,016	0,99	-2,03	0,39	3,429	0,025	0,98	-2,05	0,61
30	3,242	0,023	0,98	-2,04	0,59	3,398	0,038	0,97	-2,07	0,93
40	3,222	0,031	0,97	-2,06	0,80	3,366	0,050	0,95	-2,10	1,26
50	3,203	0,038	0,96	-2,08	1,00	3,335	0,062	0,94	-2,12	1,60
60	3,184	0,046	0,96	-2,09	1,22	3,304	0,074	0,93	-2,15	1,94
70	3,165	0,053	0,95	-2,11	1,43	3,273	0,085	0,92	-2,18	2,30
80	3,146	0,060	0,94	-2,13	1,65	3,242	0,096	0,90	-2,21	2,66
85	3,136	0,064	0,94	-2,14	1,76	3,227	0,102	0,90	-2,23	2,85
105	3,120	0,057	0,94	-2,12	1,56	3,200	0,091	0,91	-2,20	2,53
125	3,104	0,049	0,95	-2,10	1,35	3,172	0,079	0,92	-2,17	2,20
145	3,087	0,042	0,96	-2,09	1,15	3,145	0,068	0,93	-2,15	1,87
165	3,071	0,034	0,97	-2,07	0,94	3,118	0,056	0,94	-2,12	1,54
185	3,055	0,027	0,97	-2,06	0,73	3,092	0,044	0,96	-2,09	1,20
205	3,039	0,019	0,98	-2,04	0,53	3,065	0,032	0,97	-2,07	0,86
225	3,024	0,012	0,99	-2,02	0,32	3,039	0,019	0,98	-2,04	0,52
245	3,008	0,004	1,00	-2,01	0,11	3,013	0,006	0,99	-2,01	0,17
255	3,000	0,000	1,00	-2,00	0,00	3,000	0,000	1,00	-2,00	0,00

Результати розрахунків власних значень λ_1 і λ_2 , компонентів власних векторів \vec{u}_1 і \vec{u}_2 , а також кута θ між вектором \vec{u}_1 і вектором $\vec{u}_{10} = (1, 1, 1)$, що відповідає максимальному власному значенню коваріаційної матриці (6), для різних значень I_b та фіксованих значень m_1 зведемо в таблицю, наведену на с. 17.

Наведені результати показують таке:

- відношення між власними значеннями λ_2/λ_1 при зміні I_b відповідає значенню частки колірних пікселів на зображенні і становить близько одного відсотка, а отже, використання критерію $\lambda_1 > \lambda_2 > 0$; $\lambda_3 = 0$; $\lambda_2/\lambda_1 \leq 0,03$ дозволяє розрізняти зображення бланків із ахроматичним та хроматичним колірним контентом;

- кут між векторами \vec{u}_{10} та \vec{u}_1 близький до нуля, коли зображення є практично ахроматичне, а коли воно має хроматичні кольори, то θ становить близько 3° ;

- компонент u_{13} для ахроматичних кольорів дорівнює одиниці, а в разі появи хроматичних кольорів менший за одиницю. Аналогічно для ахроматичних кольорів $u_{23} = -2$, а за наявності в зображенні хроматичних кольорів $u_{23} < -2$.

Оскільки всі компоненти \vec{u}_1 додатні, то при перетворенні з простору RGB у простір головних компонентів дістаємо:

$$|\vec{u}_1| = ru_{11} + gu_{12} + bu_{13}.$$

Це відповідає перетворенню кольорів сегментів C_0 , C_1 і C_2 у градації сірого.

Висновки

Аналіз впливу колірного контенту зображень документів суворої звітності на характеристики простору Карунена–Лоева показав, що зображення з ахроматичним контентом проєктуються на вектор, який відповідає максимальному власному значенню, а простір головних компонентів є одновимірний; для зображень, що містять хроматичні кольори напису, простір головних компонентів двовимірний.

Для сегментації зображень з написом, виконаним у хроматичному кольорі, достатньо використовувати лише проєкцію зображення на власний вектор \vec{u}_2 . При цьому необхідно виділяти лише два сегменти, один з яких є рукописний текст.

Подальші дослідження мають бути спрямовані на реалізацію методу сегментації зображень документів суворої звітності та аналіз ефективності сегментації з урахуванням спотворень кольорів, що виникають при скануванні бланків.

Література

1. *Color image segmentation: Advances and prospects [Електронний ресурс] / [H. D. Cheng, X. H. Jiang, Y. Sun, Jing Li Wang].—Pattern Recognition.— 2001.— Vol. 34.— P. 2259–2281.— Режим доступу: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.119.2886&rep=rep1&type=pdf>*
2. *Castro-Bleda, M. J. Improving Offline Handwritten Text Recognition with Hybrid HMM/ANN Models / M. J. Castro-Bleda, J. Gorbe-Moya, F. Zamora-Martinez // Pattern Analysis and Machine Intelligence, IEEE Transactions on.— 2011.— Vol. 33, Issue 4.— P. 767–779.*
3. *Гладких, В. М. Класифікація цифрових зображень поштового переказу за колірним контентом / В. М. Гладких.— Наук. записки УНДІЗ.— 2011.— №2(18).*

В. Н. Гладких

ЦВЕТОВОЕ ПРОСТРАНСТВО ДЛЯ ВЫДЕЛЕНИЯ ТЕКСТА НА ИЗОБРАЖЕНИИ ДОКУМЕНТОВ СТРОГОЙ ОТЧЕТНОСТИ

Исследована модель цветового пространства преобразования Карунена–Лоева для печатных бланков документа строго учета, заполненного цветным рукописным текстом.

Ключевые слова: бланк документа; преобразование Карунена–Лоева; сегментация изображения.

V. M. Gladkykh

THE CHOICE OF COLOUR SPACE FOR THE SEGMENTATION OF DOCUMENT BLANK'S IMAGES

Investigational model of color space of Karhunen–Loeve transformation for the printing forms of documents of the severe accounting, which written in coloured handwriting text.

Keywords: form of document; Karhunen–Loeve transformation; image segmentation.