

КОМПЕНСАЦІЯ ГЕОМЕТРИЧНИХ СПОТВОРЕНЬ ЗОБРАЖЕНЬ ПРИ ОБРОБЦІ БЛАНКІВ

Кучер К.С.

Науковий керівник - докт. техн. наук, зав. каф. «Інформаційних систем», проф
Антощук С.Г.

Одне з вузьких місць при організації електронного документообігу – введення паперових документів, яке здійснюється, як правило, шляхом сканування і подальшого розпізнавання текстової інформації. Завдання розпізнавання рукописного тексту в структурованих документах (відомостях, бланках, анкетах) є актуальним, оскільки полегшує введення рукописних даних у комп'ютер.

При розпізнаванні відсканованих документів на першому етапі роботи проводиться попередня обробка, завданням якої є покращення якості вихідного напівтонового зображення. Її необхідність викликана тим, що вихідні відскановані зображення, як правило, мають спотворення, такі, як поява темних областей на краях внаслідок нерівномірного освітлення та скошування зображення. [1]

Видалення масивних темних областей на краях зображення вирішується, як правило, засобами порогової обробки. При дослідженні типових вихідних зображень було визначено, що такі області можуть займати не більше 25% від кожного краю зображення. Експериментальним шляхом було визначено порогове значення як середню інтенсивність, яка відповідає темній області.

Задача компенсації скошування зображення (*skew estimation*) є більш складною. Аналіз показав, що існуючі методи компенсації поділяються на дві групи: методи, що оцінюють кут нахилу за прямими лініями, знайденими на зображенні, та методи, що оцінюють кут нахилу за характеристиками зон інтересу на зображенні документа [2]. Методи обох груп було реалізовано та порівняно за якістю.

Розглянемо першу групу. Найпростіший метод обчислення нахилу базується на побудові обмежувального прямокутника, сторони якого не є паралельними до осей (*non-axis aligned bounding box*), і використовує координати двох верхніх крайніх точок заповнення зображення. Кут нахилу зображення φ обчислюється за формулою:

$$\varphi = \arctg \frac{y_2 - y_1}{x_2 - x_1},$$

де $(x_1; y_1)$ та $(x_2; y_2)$ – координати лівої та правої верхніх точок зон інтересу на зображенні. Для пошуку точок використовується сканування прямокутними областями від верхніх кутів зображення. Недоліки такого методу проявляються навіть на зображеннях, на яких верхні крайні точки відповідають одній лінії тексту: якщо буде знайдено гліфи різної висоти чи якщо при порівнянні значення пікселя із граничним у наслідок напівтонової природи зображення буде відкинута справжню крайню точку гліфа, то кут буде визначено невірно. Серед переваг цього методу – швидкість роботи: для аналізу напівтонового зображення 867 на 630 пікселів у середовищі *Scilab* витрачається 0.09 секунди.

Також до методів, що використовують виділення на зображенні прямих ліній, відносяться такі, що базуються на використанні перетворення Хафа (*Hough Transform*): кожна дискретна точка зображення тексту перетворюється у криву на площині у полярній системі координат. Ця площина розбивається на комірки. Якщо будь-який фрагмент кривої перетинає комірку, то вагове значення комірки збільшується. У результаті можна прослідити, де перетинаються криві, що відповідають точкам вихідного зображення, тобто знайти точки, що належать одній прямій. Головні недоліки цього перетворення при компенсації спотворення зображення – складність інтерпретації результатів перетворення (що пояснюється великою кількістю точок зон, між якими встановлюються зв'язки) та низька швидкість роботи перетворення (на саме перетворення Хафа у середовищі *Scilab* для напівтонового зображення 867 на 630 пікселів витрачається 25.58 секунд).

Друга група методів використовує статистичну інформацію про заповнення зображення. Замість побудови обмежувального прямокутника зона інтересу описується еліпсом розсіювання [3], у якості значення орієнтації якого використовують напрямок власного вектору матриці коваріації координат точок заповнення (на зображенні бланка – це темні пікселі тексту). Кут нахилу φ обчислюється за формулою:

$$\varphi = \arctg \left(\frac{D(x) - \frac{D(x) - D(y)}{2} - \sqrt{\left(\frac{D(x) - D(y)}{2}\right)^2 - D(x)D(y) + C^2(x, y)}}{-C(x, y)} \right),$$

де $D(x), D(y)$ – дисперсія горизонтальних та вертикальних координат точок зон інтересу відповідно, $C(x, y)$ – коваріація координат точок зон інтересу.

Цей метод демонструє достатню якість результату та швидкість (для аналізу напівтонового зображення 867 на 630 пікселів у середовищі *Scilab* витрачається 0.1 секунди), тому його було обрано для компенсації скошування зображення. Для подальшого

покращення якості було вирішено використовувати задану користувачем інформацію про розміщення зон інтересу на зображенні, що використовується при розпізнаванні інформації.

Реалізація даного метода дозволяє компенсувати спотворення зображення скошуванням для задачі обробки бланків.

СПИСОК ЛІТЕРАТУРИ

1. Горский Н., Анисимов В. Распознавание рукописного текста // СПб: Политехника, 1997. – 256 с.
2. R. Cattoni, T. Coianiz, S. Messelodi, C. M. Modena. Geometric Layout Analysis Techniques for Document Image Understanding. – ITC-IRST, Via Sommarive, I-38050 Povo, Trento, Italy – 1998.
3. В.Т. Фисенко, Т.Ю. Фисенко, Компьютерная обработка и распознавание изображений: учеб. пособие. – СПб: СПбГУ ИТМО, 2008. – 192 с.