

## ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ВІДНОВЛЕННЯ СПОТВОРЕНИХ ЗОБРАЖЕНЬ ТЕКСТОВИХ ДОКУМЕНТІВ

Описано метод відновлення спотворених зображень текстових документів. Виправлення спотворень відбувається у три етапи. Перший етап (грубе вирівнювання) спрямований на усунення сильних деформацій документа та отримання першого результату корекції. На другому етапі відбувається геометрична корекція локальних деформацій зображення документа на рівні слів та отримання виправленого зображення із горизонтально рівними рядками тексту. Третій етап методу спрямований на усунення перспективних спотворень шляхом виявлення горизонтальної та вертикальної точок сходження та отримання виправленого зображення текстового документу без будь яких деформацій.

**Ключові слова:** спотворені зображення, зображення текстового документа, відновлення деформованого зображення

The method of restoring the distorted images of text documents. Correcting distortion occurs in three stages. The first stage (coarse alignment) is aimed at addressing the severe strain document and obtaining first result correction. In the second stage, the geometric correction of image deformations local document with words and getting the corrected image with equal horizontal lines of text. The third stage of the method aims to eliminate perspective distortion by horizontal and vertical identify points of convergence and obtaining a corrected image text document without any deformation.

**Keywords:** distorted image, images, text documents, restore the deformed image.

### Вступ

Існуючі методи корекції спотворених зображень текстових документів [1-8] поділяються на два основних види, в залежності від підходу до розв'язання проблеми: (а) методи, засновані на тривимірній реконструкції документів [6-8] та (б) методи, засновані тільки на двовимірній обробці документу [1-5]. Методи, що відносяться до першої категорії, є досить ефективними, проте обмежують середовище використання цифрових фотокамер і потребують додаткового обладнання (наприклад лазерних сканерів, стерео-камер, структурованих джерел світла) або ж накладаються певні початкові обмеження, щоб відтворити форму поверхні сторінки. Методи, віднесені до другої категорії, передбачають використання камер у неконтрольованому середовищі і вимагають лише інформацію про

---

<sup>1</sup> Українська академія друкарства

<sup>2</sup> Uniwersytet Warmińsko-Mazurski w Olsztynie

розташування рядків без побудови форми поверхні сторінки. Більшість методів, що відносяться до другої категорії побудовані на основі сегментації ліній текстових рядків. Проте, сегментація текстових рядків спотворених зображень є складним процесом, що часто призводить до помилок.

Представлений метод заснований на двовимірній обробці документів і не потребує додаткового обладнання для відновлення форми поверхні сторінки. Щоб уникнути помилок в процесі сегментації текстових рядків, спочатку здійснюється грубе виправлення спотворень, а далі більш точне виправлення геометричних спотворень шляхом виділення слів, їх поворотом та переміщенням та перспективних спотворень на основі виявлення точок сходження.

### **1. Алгоритм відновлення спотворених зображень**

Перед початком застосування даного методу здійснюється попередня обробка зображення – бінаризація, видалення шуму та нетекстових компонент. Для бінаризації використовується метод, на основі глобального порогового методу Отсу та додаткова обробка для корекції помилок процесу бінаризації і видалення шуму. Процес попередньої обробки більш детально описаний в [9].

Далі, вже на чорно-білому зображенні, метод використовує три основних етапи:

- корекція форми поверхні сторінки – на основі побудови моделі перетворення, яка б відображала проекцію вигнутої поверхні у прямокутну область
- корекція геометричних спотворень - шляхом виділення границь слів і переміщаються зображень у відповідну лінію рядка
- корекція перспективних спотворень зображення на основі виявлення горизонтальної та вертикальної точки сходження

Загальна схема даного методу зображена на рис.1.

### **2. Корекція форми поверхні сторінки**

Потрібно побудувати модель перетворення, яка б відображала проекцію вигнутої поверхні у двовимірну прямокутну область. Для знаходження проекції вигнутої поверхні потрібно оцінити текстові межі документа. Ліву та праву межу тексту оцінюємо відрізками, методом найменших квадратів [10] на основі всіх виявлених найбільш лівих/правих точок, за винятком тих точок рядків тексту, що не починаються з початку документа (заголовки, підзаголовки). Для оцінки верхньої та нижньої межі тексту поліномом третього степеня теж використовується метод найменших квадратів.

Таким чином, ми отримуємо два відрізки:  $AD$ , що відповідає лівій межі тексту

$$y = a_l x + b_l \quad (1)$$

та  $BC$ , що відповідає правій межі тексту:

$$y = a_r x + b_r \quad (2)$$

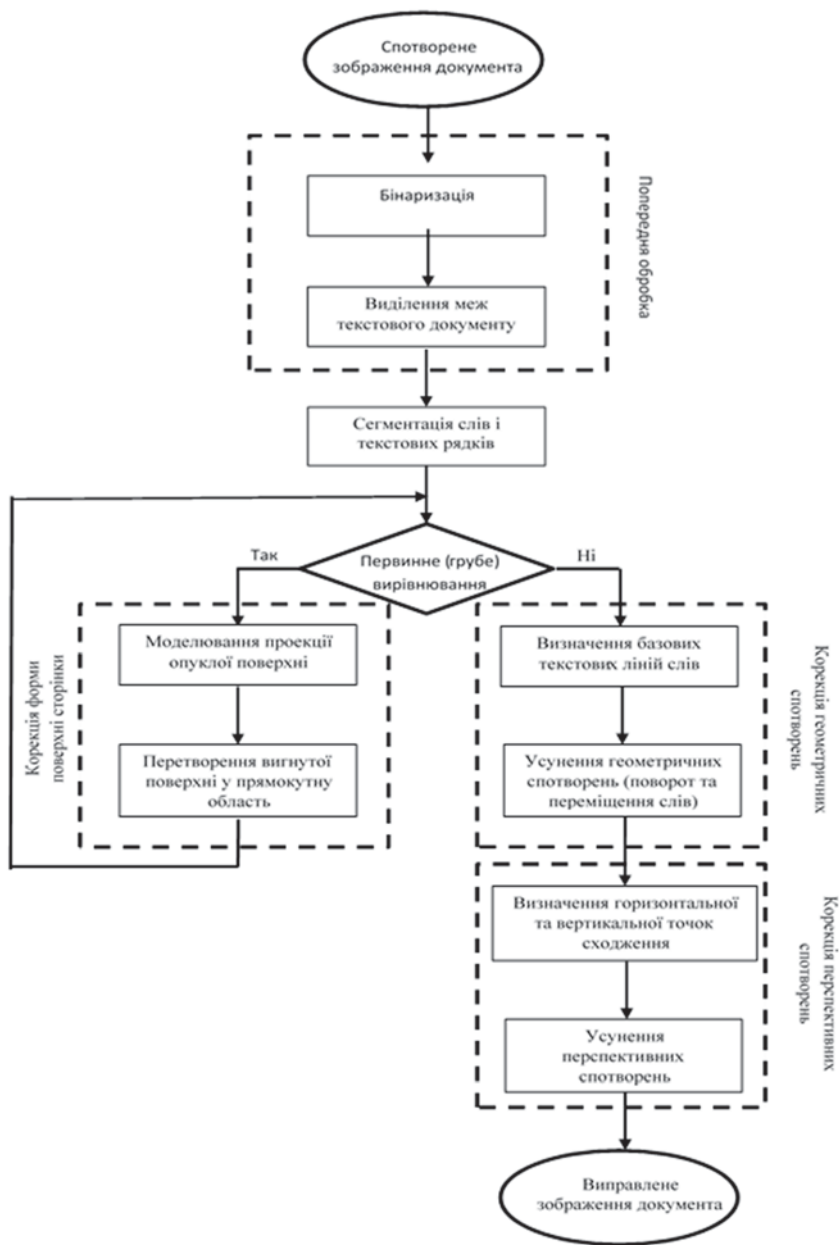


Рис.1: Інформаційна технологія відновлення спотворених зображень текстових документів

Рівняння кривої верхньої межі тексту  $AB$  буде мати вигляд:

$$y = a_{u1}x^3 + a_{u2}x^2 + a_{u3}x + a_{u4} \quad (3)$$

Аналогічно, крива  $BC$ :

$$y = a_{l1}x^3 + a_{l2}x^2 + a_{l3}x + a_{l4} \quad (4)$$

Потрібно побудувати перетворення, яке б відображало проекцію зігнутої поверхні, обмежену кривими  $AB, DC$  і прямими  $AD, BC$  у двовимірну прямокутну область. Нехай  $A'(x'_1, y'_1)$ ,  $B'(x'_2, y'_2)$ ,  $C'(x'_3, y'_3)$ ,  $D'(x'_4, y'_4)$  – кутові точки прямокутної області (рис.2),  $|\widehat{AB}|$  – довжина дуги між точками  $A$  і  $B$ ,  $|AB|$  – евклідова відстань між точками  $A$  і  $B$ .

Ширина  $W$  прямокутної області визначається так:

$$W = \min(|\widehat{AB}|, |\widehat{DC}|) \quad (5)$$

Висота  $H$  прямокутної області дорівнює:

$$H = \min(|AD|, |BC|) \quad (6)$$

У нашому прикладі (рис. 2)  $W = |\widehat{AB}|, H = |AD|$ .

Кутові точки прямокутної області обчислюються наступним чином:

Побудуємо функцію  $\Phi$  для створення відповідності між кривими  $AB$  і  $DC$ .

$$\Phi(E(x_u, y_u)) = G(x_l, y_l), \text{ якщо } \frac{|\widehat{AE}|}{|\widehat{AB}|} = \frac{|\widehat{DG}|}{|\widehat{DC}|}, \quad (8)$$

де  $E(x_u, y_u)$  – точка на кривій  $AB$ ,  $G(x_l, y_l)$  – точка на кривій  $DC$ .

Для всіх точок з проекції зігнутої поверхні визначаємо нове положення. Нехай  $O(x, y)$  – точка на проекції зігнутої поверхні. Наше завдання визначити нове положення  $O'(x', y')$  точки  $O(x, y)$  (рис.2).

Спочатку визначимо пряму  $EG$ , яка задовольняє наступним умовам:

1. Перетинає точки  $E(x_u, y_u)$ ,  $G(x_l, y_l)$ , що лежать на кривих  $AB$  і  $DC$  відповідно.

2.  $\Phi(E(x_u, y_u)) = G(x_l, y_l)$

3. Точка  $O(x, y)$  належить прямій  $EG$ .

Далі обчислюємо положення  $O'(x', y')$ :

$$x' = x'_1 + |A'Z| \quad (9)$$

$$y' = y'_1 + |A'H| \quad (10)$$

де  $H$  – це точка  $H(x'_1, y')$ ,  $Z$  – це точка  $Z(x', y'_1)$ . Довжини відрізків  $|A'Z|$ ,  $|A'H|$  обчислюються наступним чином:

$$\frac{|\widehat{AB}|}{|\widehat{AE}|} = \frac{W}{|A'Z|} \Rightarrow |A'Z| = \frac{W}{|\widehat{AB}|} |\widehat{AE}| \quad (11)$$

$$\frac{|EG|}{|EO|} = \frac{H}{|A'H|} \Rightarrow |A'H| = \frac{H}{|EG|} |EO| \quad (12)$$

Дану послідовність дій повторюємо для всіх точок з проекції зігнутої поверхні. Якщо точка знаходиться поза областю, вона приймає перетворення найближчої точки.

$$\left. \begin{aligned} x'_1 &= x_1, & y'_1 &= y_1 \\ x'_2 &= x'_1 + W, & y'_2 &= y'_1 \\ x'_3 &= x'_2, & y'_3 &= y'_2 + H \\ x'_4 &= x'_1, & y'_4 &= y'_3 \end{aligned} \right\} \quad (7)$$

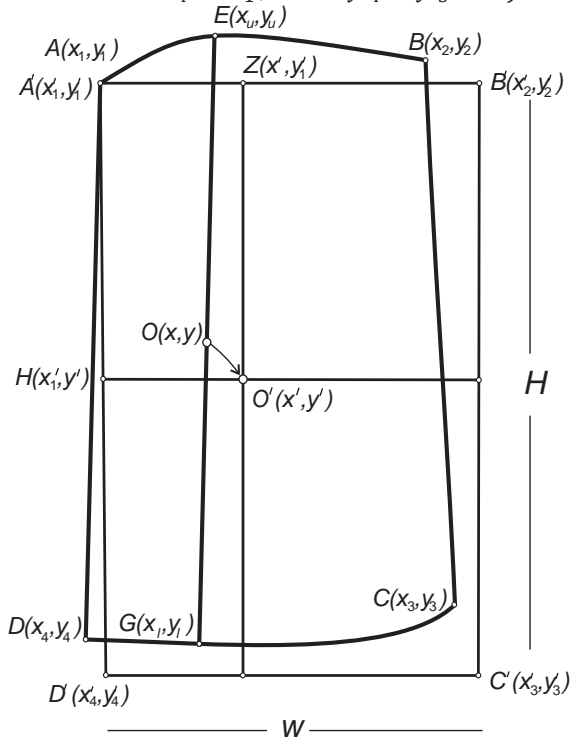


Рис.2: Проекція вигнутої поверхні

### 3. Корекція геометричних спотворень

Спочатку виділяються всі текстові рядки і слова на зображенні. Нехай  $L_C$  - кількість рядків, а  $H$  – середня висота символу. Усі слова виділяються за допомогою належного згладжування зображення; потім горизонтально сусідні слова послідовно з'єднуються з метою визначення текстових рядків. Цей процес детальніше описаний в роботі [11], блок-схема зображена на рис.3.

У кінці цієї процедури, кожен рядок тексту розглядається як  $L_i$ , а кожне слово  $j$ , що належить лінії  $L_i$  як  $W_{ij}$ . Останнім завданням цього етапу є: об'єднання перших двох слів (ліворуч) кожного текстового рядка, якщо ширина першого слова є малою. Ці умови є необхідними, оскільки для коротких слів важко виявити нахил слова і тому неможливо перейти до

наступного кроку. З іншого боку, важливо мати точне визначення нахилу для першого слова з лівого боку, яке дає напрямок вирівнювання для цілого рядка тексту.

Для наближення нижньої і верхньої базової лінії слова використовується лінійна регресія, яка будується на множині точок, які є найнижчими (для нижньої межі слова) чи верхніми (для верхньої межі) чорними пікселями для кожного зображення у колонці. Після цієї процедури, верхня базова лінія слова  $W_{ij}$  визначається наступним чином:

$$y = a_{ij}x + b_{ij}, \quad (13)$$

нижня базова лінія слова  $W_{ij}$  визначається як

$$y = a'_{ij}x + b'_{ij} \quad (14)$$

Нахил кожного слова отримується з відповідних нахилів базових ліній. Нахил верхньої і нижньої базової лінії  $\theta_{ij}^u$  і  $\theta_{ij}^l$  слова  $W_{ij}$  визначається як

$$\theta_{ij}^u = \arctan(a_{ij}), \quad \theta_{ij}^l = \arctan(a'_{ij}) \quad (15)$$

Оскільки найбільш характерним є, як правило, менший абсолютний нахил, то нахил самого лівого слова може бути визначений як:

$$\theta_{i0} = \begin{cases} \theta_{i0}^u, & \text{якщо } |\theta_{i0}^u| < |\theta_{i0}^l| \\ \theta_{i0}^l, & \text{в іншому випадку} \end{cases} \quad (16)$$

Нахил усіх інших слів є одним з найближчих значень до нахилу попереднього слова:

$$\theta_{ij} = \begin{cases} \theta_{ij}^u, & \text{якщо } |\theta_{ij}^u - \theta_{ij-1}| < |\theta_{ij}^l - \theta_{ij-1}| \\ \theta_{ij}^l, & \text{в іншому випадку} \end{cases}, \quad (17)$$

де  $j > 0$ .

Поворот слова  $W_{ij}(x, y)$  обчислюється наступним чином:

$$\begin{cases} y^r = (x - x_{ij}^{min}) * \sin(-\theta_{ij}) + y * \cos(\theta_{ij}) \\ x^r = x \end{cases} \quad (18)$$

де  $W_{ij}^r(x^r, y^r)$  – повернуте слово,  $x_{ij}^{min}$  – ліва обмежувальна координата слова  $W_{ij}$ .

Після обертання слів, всі слова кожного текстового рядка, окрім самого лівого, повинні бути переміщені вертикально в цілях відновлення горизонтального вирівнювання. Обертання і переміщення слова  $W_{ij}(x, y)$  робиться таким чином:

$$\begin{cases} y^{rs} = y^r + d_{ij} \\ x^{rs} = x \end{cases} \quad (19)$$

де  $W_{ij}^{rs}(x^{rs}, y^{rs})$  – повернуте і переміщене слово,  $y^r$  визначається з системи (19),  $d_{ij}$  відповідає вертикальному зсуву слова і визначається за наступною формулою:

$$d_{ij} = \begin{cases} y_{i0}^{ru} - y_{ij}^{ru}, & \text{якщо } |\theta_{ij}^u - \theta_{ij-1}| < |\theta_{i0}^l - \theta_{ij-1}| \\ y_{i0}^{rl} - y_{ij}^{rl}, & \text{в іншому випадку} \end{cases}, \quad (20)$$

де  $y_{ij}^{ru} = (a_{ij}x_{ij}^{min} + b_{ij}) * \cos(\theta_{ij})$  і  $y_{ij}^{rl} = (a'_{ij}x_{ij}^{min} + b'_{ij}) * \cos(\theta_{ij})$ .

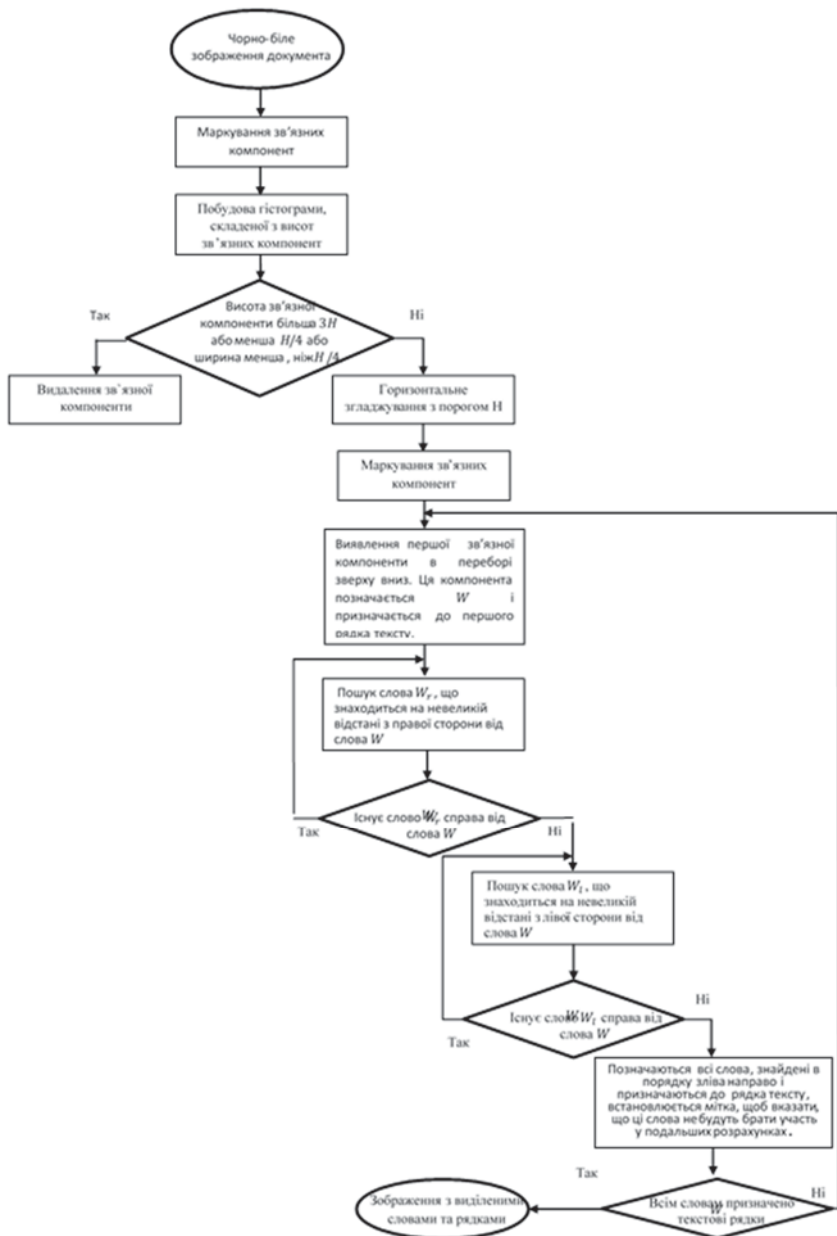


Рис.3: Алгоритм виділення слів та рядків на зображенні

Причиною наявності двох варіантів переміщення є те, що кожне слово може бути посунуте або до його нижньої базової лінії або до схилу верхньої базової лінії. Отже, воно має бути зміщеним так, щоб його нижня чи верхня базова лінія сумістилась з нижньою чи верхньою базовою лінією лівого слова рядка тексту.

Далі додаються всі нетекстові компоненти, які були раніше вилучені. Для цього кожен піксель приймає значення коефіцієнтів трансформації  $(x^{rs}, y^{rs})$  найближчого пікселя. Потім застосуємо перетворення для кожної компоненти, яке використовує в якості факторів перетворення середнє значення складових факторів пікселів.

#### 4. Корекція перспективних спотворень

У випадку сильних спотворень (рис. 4), чи коли на зображенні має місце комбінація декількох типів спотворень, навіть після вирівнювання рядків залишаються вертикальні перекошування символів. У таких випадках необхідно аналізувати вертикальні кути нахилів символів. Один з методів, описаний в [12].

Припускається, що на вихідному спотвореному зображенні всі текстові рядки вирівняні за допомогою способу, описаного в пунктах 2,3. Запропонований алгоритм побудований на наступних етапах:

- 1) Визначення рівня прямої кривої
- 2) Визначення горизонтальних базових текстових ліній / визначення вертикальних штрихів символів
- 3) Визначення горизонтальної / вертикальної точки сходження
- 4) Випрямлення перспективних спотворень зображення

У графічній перспективі точка сходження — це точка, де сходяться паралельні лінії.

Точка сходження (горизонтальна або вертикальна) у двовимірному просторі може бути описана як  $(v_x, v_y)$ , де  $v_x$  і  $v_y$  координати осей  $x$  та  $y$  відповідно у двовимірному евклідовому просторі

$$R^2 = \{(x, y) | -\infty < x < +\infty, -\infty < y < +\infty\}, \quad (21)$$

точка  $(0,0)$  – центр зображення.

Знайти точку сходження означає визначити її з деякого процесу оптимізації на площині зображення. Тобто,

$$(v_x, v_y) = \arg \max_{(x,y) \in R^2} f(x, y) \quad (22)$$

де  $f(x, y)$  – функція оптимізації.

Якщо вибирати точки сходження із всього простору  $R^2$ , такий процес займе дуже багато часу для обчислень. Тому потрібно зменшити множину можливих значень, об'єднавши в групи точки перетину ліній як можливі точки сходження.

Спочатку знаходять всі прямі і умовно прямі лінії. Потім розраховується точка перетину для кожної пари ліній. Набір отриманих точок перетину:

$$Set(Pt) = \{(x_1, y_1), (x_2, y_2), \dots, (x_{N_{Pt}}, y_{N_{Pt}})\}, \quad (23)$$



де  $N_{pt}$  – кількість точок перетину. Ці точки розділяємо на кілька груп за допомогою алгоритму кластеризації. Таким чином, ми отримуємо нову множину  $S$

$$S = \{S_i | S_i \subset R^2, i = 1, \dots, N\} \quad (24)$$

Вибирається центральна точка перетину кожної групи і розглядається як єдиний представник даної підобласті. Множина центральних точок перетину визначається наступним чином:

$$C = \{c_i | c_i \in S_i, i = 1, \dots, N\}. \quad (25)$$

Замість того, щоб шукати точку сходження на всьому просторі  $R^2$ , ми здійснюємо пошук в  $C$  для збільшення швидкодії алгоритму. Оскільки множина точок з  $C$  є достатньо показовою, пошук максимуму в  $C$  є хорошим наближенням до глобального максимуму.

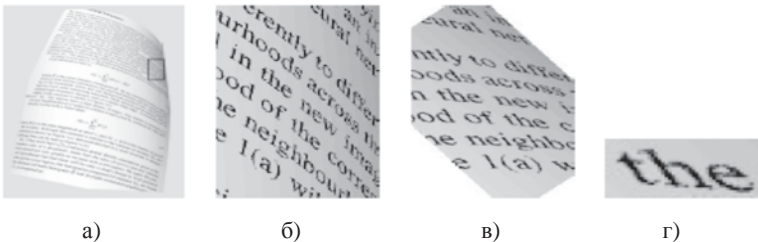


Рис. 4 Приклад сфотографованого зображення документу: (а) фотографія вигнутої поверхні документу; (б) частина збільшеного зображення; (в) вирівняні рядки до горизонтального вигляду; (г) на рівні символів (навіть після вирівнювання рядків, спотворені символи всеодно важко розпізнати)

Отже, точка сходження виражається наступним рівнянням:

$$(v_x, v_y) = \arg \max_{(x,y) \in C} f(x, y) \quad (26)$$

Враховуючи горизонтальні і вертикальні точки сходження, можна відновити документи з перспективними спотвореннями. Рівняння перетворення визначається наступним чином:

$$\begin{pmatrix} x_d \\ y_d \\ 1 \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix} \begin{pmatrix} x_u \\ y_u \\ 1 \end{pmatrix}, \quad (27)$$

де  $(x_u, y_u)$  – координати виправленого зображення,  $(x_d, y_d)$  – координати спотвореного зображення.

Враховуючи горизонтальні і вертикальні точки сходження на площині зображення, ми можемо обчислити опуклий чотирикутник на площині зображення, який відповідає прямокутнику на випрамленій площині зображення. Метод для виявлення таких опуклих чотирикутників зображено на рис.5.

Співвідношення сторін результуючого зображення визначається наступним чином. Середня довжина верхньої та нижньої сторін

чотирикутника відповідає ширині результуючого зображення, а середня довжина лівої і правої сторін чотирикутника відповідає висоті результуючого зображення.

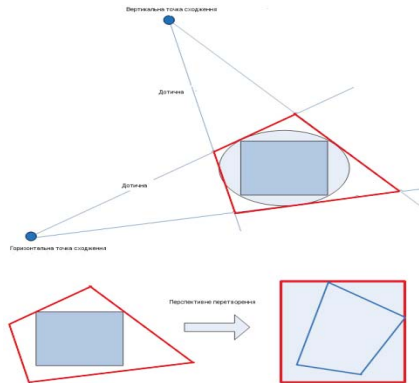


Рис.5 Перспективне перетворення

## Висновки

У статті описано метод відновлення спотворених зображень, що характеризуються нерівною поверхнею, геометричними та перспективними спотвореннями; побудовано загальну схему методу. Метод поділено на 3 етапи, на кожному з яких частково виправляються спотворення, а отже і покращується якість розпізнавання зображення. На першому етапі усуваються сильні деформації зображення, тому проведена на другому етапі сегментація рядків і слів є більш точною. Даний спосіб відновлення зображень є ефективним при застосуванні до сильно спотворених зображень текстових документів.

1. Масалович А.А. Использование патча Безье для аппроксимации искажения текстовых документов / А.А. Масалович, Л.М. Местецкий // Труды 17-ой Международной Конференции по Компьютерной Графике и Зрению (Графикон-2007). – Москва. – 2007. - 239-243 ст. - URL: [http://www.machinelearning.ru/wiki/images/a/a2/Gr\\_2007\\_Masalovich.pdf](http://www.machinelearning.ru/wiki/images/a/a2/Gr_2007_Masalovich.pdf)
2. Bin Fu, Minghui Wu, Rongfeng Li, Wenxin Li, Zhuoqun Xu, Chunxu Yang, "A model based book dewarping method using text line detection", Proceedings of the Second International Workshop on Camera-Based Document Analysis and Recognition (CBDAR-2007), 2007, Curitiba, Brazil, pp. 63-70. - URL: <http://imlab.jp/cbdar2007/proceedings/papers/P1.pdf>
3. Minghui Wu, Rongfeng Li, Bin Fu, Wenxin Li, Zhuoqun Xu, "A model based book dewarping method to handle 2D images captured by a digital camera", Proceedings of the 9-th International Conference on Document Analysis and Recognition (ICDAR-2007), 2007, Curitiba, Brazil, pp. 158-162.
4. Li Zhang, Chew Lim Tan, "Warped Image Restoration with Application to Digital Libraries", Proceedings of the 8-th International Conference on

Document Analysis and Recognition (ICDAR-2005), 2005, Seoul, South Korea, pp. 192-196.124

5. *Xu-Cheng Yin, Jun Sun, Satoshi Naoi*, "Perspective rectification for mobile phone camera-based documents using a hybrid approach to vanishing point detection", Proceedings of the Second International Workshop on Camera-Based Document Analysis and Recognition (CBDAR-2007), 2007, Curitiba, Brazil, pp. 37-44.

A. *Ulges, C. Lampert, and T. M. Breuel*. Document capture using stereo vision. In Proceedings of the ACM Symposium on Document Engineering, pages 198–200. ACM, 2004.

6. *Yamashita, A. Kawarago, T. Kaneko, and K.T. Miura*. Shape reconstruction and image restoration for non-flat surfaces of documents with a stereo vision system. In Proceedings of 17th International Conference on Pattern Recognition (ICPR2004), Vol.1, pages 482–485, 2004.

7. *M.S. Brown and W.B. Seales*. Document restoration using 3d shape: A general deskewing algorithm for arbitrarily warped documents. In International Conference on Computer Vision (ICCV01), volume 2, pages 367–374, July 2001.

8. *Кульчицька І.О., Тимченко О.В.* Особливості алгоритмів бінаризації зображень документів // Зб. наук. пр. ІПМЕ НАН України. – Вип.68. – К.: 2013. – С.141-149.

9. *Линник Ю. В.* Метод наименьших квадратов и основы математико-статистической теории обработки наблюдений / Юрий Владимирович Линник. – Государственное издательство Физико-математической литературы. – 1958

10. *Тимченко О.В., Кульчицька І.О., Тимченко О.О.* Відновлення геометрії довільно спотворених зображень документів шляхом сегментації // Моделювання та інформаційні технології. Зб. наук. пр. ІПМЕ НАН України. – Вип.70. – К.: 2013. – С.171-176.

11. *Кульчицька І.О.* Метод корекції перспективних спотворень на зображеннях текстових документів // Моделювання та інформаційні технології. Зб. наук. пр. ІПМЕ НАН України. – Вип.71. – К.: 2014. – С. 153-159

*Поступила 7.9.2015р.*

УДК 004.9

Б.В.Дурняк, М.М. Кляп, УАД, м.Львів

## **ОРГАНІЗАЦІЯ СИСТЕМИ ПРОГНОЗУВАННЯ, ПРИ ВИНИКНЕННІ НЕОЧІКУВАНИХ НЕГАТИВНИХ ФАКТОРІВ**

### **Вступ**

В рамках загальної організації системи управління *DTP*, при виникненні негативних зовнішніх факторів, що орієнтовані на реалізацію впливу на технологічний процес, необхідно більш детально зупинитися на компоненті, що представляє собою модель зовнішніх процесів  $Z P_i$ . Цю компоненту необхідно розглянути з точки зору наступних аспектів її функціонування: