

УДК 004.9

ДОСЛІДЖЕННЯ ТА ОЦІНКА МЕТОДІВ КОРЕКЦІЇ ГЕОМЕТРИЧНИХ СПОТВОРЕНЬ ЗОБРАЖЕНЬ ТЕКСТОВИХ ДОКУМЕНТІВ

О. В. Тимченко^{1,2}, І. О. Кульчицька¹, Р. О. Кульчицький¹, О. О.Тимченко¹

¹Українська академія друкарства,
вул. Під Голоском, 19, Львів, 79020, Україна

²Вармінсько-Мазурський університет,
вул. М. Очаповського, 2, Ольштин, 10-719, Польща

Проаналізовано відомі методи корекції геометричних спотворень зображень текстових документів та недоліки кожної групи цих методів. Розроблено новий метод корекції спотворень, що не залежить від типу спотворення та може застосовуватися до зображень із комбінацією декількох типів спотворень. Розглянуто особливості оцінки алгоритмів корекції спотворень зображень за допомогою систем OCR. Проведено експериментальне дослідження, яке показало, що застосування розроблених методів корекції спотворень на етапі попередньої обробки перед розпізнаванням тексту дає можливість істотно підвищити якість розпізнавання. Експериментальні дослідження підтвердили, що розроблені методи корекції геометричних і перспективних спотворень забезпечують вищу якість попередньої обробки, ніж комерційна програма BookRestorer.

Ключові слова: спотворені зображення, системи OCR, зображення текстового документу, якість розпізнавання, корекція спотворення, попередня обробка зображень.

Постановка проблеми. На зображеннях текстових документів, отриманих за допомогою фотоапарата, камери мобільного телефону чи сканера, дуже часто з'являються різного типу спотворення, пов'язані з нерівністю поверхні вихідного документа, нерівномірним освітленням, нерівним положенням документа у сканері відносно сторін, а також спотворення можуть виникати, коли площина фотоапарата не паралельна площині документа, коли сканується товста книга (утворюється затемнена ділянка в місці згину книги) і т. д. Такі зображення є незручними для читання та значно погіршують якість розпізнавання тексту системами OCR. Тому в таких випадках треба провести попередню обробку зображення для усунення його дефектів.

Корекція геометричних і перспективних спотворень у системах OCR залишається актуальною та має велике практичне значення, оскільки майже всі системи оптичного розпізнавання тексту розраховані на те, що рядки зображення тексту, який розпізнається, будуть прямими та горизонтальними. Відомі методи корекції геометричних спотворень можуть незначно виправити лише один окремий тип

спотворення (геометричний, перспективний чи ін.) і не можуть застосовуватися до зображень із комбінацією декількох типів спотворень.

Аналіз останніх досліджень та публікацій. Останнім часом були запропоновані та описані в [1–15] нові методи вирішення проблеми корекції геометричних спотворень, які можна застосовувати до зображень текстових документів. Ці методи можуть бути поділені на дві категорії залежно від підходу до розв'язання проблеми:

- 1) методи виявлення і реконструкції поверхні сторінки [1–8];
- 2) методи аналізу та випрямлення ліній рядків тексту [9–15].

Методи на основі тривимірної реконструкції документа, що належать до першої категорії, поділяються на дві підкатегорії. Перші відновлюють тривимірну структуру документа на основі додаткової інформації — за допомогою спеціального обладнання, такого як лазерний сканер [1], стереокамери [2, 3] або тривимірні сканери (3D-сканери), що працюють за технологією структурованого світла [4]. Використання спеціального обладнання зменшує привабливість застосування цих методів.

Методи, які належать до другої підкатегорії, ґрунтуються тільки на інформації, що міститься в оцифрованому документі. Сао та ін. [5], щоб виправити деформації документа, використовують циліндричну модель. Параметри моделі побудовані на скелеті горизонтальних текстових ліній документа. Проте цей метод не є універсальним, тому що спотворення зображень текстових документів не завжди можна наблизити циліндричною моделлю. У праці [6] Liang та ін. розраховують тривимірну структуру документа безпосередньо з аналізу моделі форми сторінки і властивостей текстури потоку. Метод потребує достатнього обсягу тексту в документі. Тап і ін. [7] та Zhang і ін. [8], щоб відновити тривимірну структуру документа, користуються інформацією про затінені ділянки, які утворюються на вигнутих документах. Однак часто немає інформації про рівень освітленості документів, а тому застосування такого підходу є обмежене.

До другого класу методів (на основі двовимірної обробки поверхні документа) належать методи, засновані тільки на інформації, яка міститься в оцифрованих документах, і працюють вони лише в двовимірному просторі для корекції деформацій. Більшість методів цього класу побудовані на основі визначення вигнутих ліній тексту спотворених документів. Цей процес досить складний і часто може привести до помилкових результатів.

Методи, які належать до першої категорії, є досить ефективними, проте обмежують середовище застосування цифрових фотокамер і потребують додаткового обладнання (наприклад, лазерних сканерів, стереокамер, структурованих джерел світла) або ж накладаються певні початкові обмеження, щоб відтворити форму поверхні сторінки. Методи, що належать до другої категорії, передбачають використання камер у неконтрольованому середовищі та потребують лише інформації про розміщення рядків без побудови форми поверхні сторінки. Методи другої категорії ґрунтуються на визначених моделях спотворення документа і залежать від типу спотворення. Спочатку виділяються текстові рядки, і на основі цієї інформації будується функція спотворення зображення. Потім зображення

випрямляється. Процес виділення рядків на зображенні досить складний і часто може привести до помилкових результатів.

Отже, всі проаналізовані методи мають відповідні недоліки, не є універсальними і можуть бути використані лише тоді, коли на зображенні існує тільки один з усіх видів спотворень. Тому корекція складних геометричних та перспективних спотворень текстових документів залишається актуальним завданням.

Мета статті — проаналізувати відомі методи корекції геометричних та перспективних спотворень зображень текстового документа і на його основі розробити метод усунення геометричних та перспективних спотворень, який не залежить від типу спотворення, оцінити його ефективність за допомогою експериментальних досліджень для порівняння розробленого методу з наявними комерційними програмами.

Виклад основного матеріалу дослідження. Розроблений метод належить до другої групи методів корекції спотворень, заснований на двовимірній обробці документів і не потребує додаткового обладнання для відновлення форми поверхні сторінки. Додатково, на відміну від відомих методів, не ґрунтується на строгих припущеннях для моделювання форми поверхні сторінки. Таким чином, врахувавши недоліки першої та другої груп методів корекції спотворених зображень, розроблений метод не залежить від типу спотворень і може бути застосований до зображень із комбінацією декількох типів спотворень. Щоб уникнути помилок у процесі сегментації текстових рядків, спочатку здійснюється грубе виправлення спотворень, а відтак точніше виправлення геометричних спотворень за допомогою виділення слів, їх повороту та переміщення і перспективних спотворень на основі виявлення точок сходження.

Автори розробили новий метод корекції спотворених зображень, описаний в [16], а в пропонованій праці досліджуємо якість його функціонування.

Основні етапи розробленого методу такі. Спочатку здійснюється попередня обробка зображення — бінаризація, видалення шуму та нетекстових компонент. Для бінаризації доцільно застосувати ефективний метод на основі глобального порогового методу Отсу та виконати додаткову обробку для корекції помилок процесу бінаризації і видалення шуму [17]. Далі, вже на чорно-білому зображенні, в методі використовуємо три основні етапи:

- 1) корекція форми поверхні сторінки — на основі побудови моделі перетворення, яка відображає проекцію вигнутої поверхні у прямокутну ділянку;
- 2) корекція геометричних спотворень — за допомогою виділення границь слів і переміщення зображень слів у відповідну лінію рядка;
- 3) корекція перспективних спотворень зображення на основі виявлення горизонтальної та вертикальної точки сходження.

Загальна блок-схема цього методу зображена на рис. 1.

Як показано на блок-схемі, сегментація текстових рядків проводиться на кожному етапі, тому точність процесу сегментації на кожному етапі методу покращується. Спочатку сегментація застосовується до первинного спотвореного документа, відтак до результату після грубого виправлення документа і, нарешті, до зображення, вільного від геометричних спотворень.

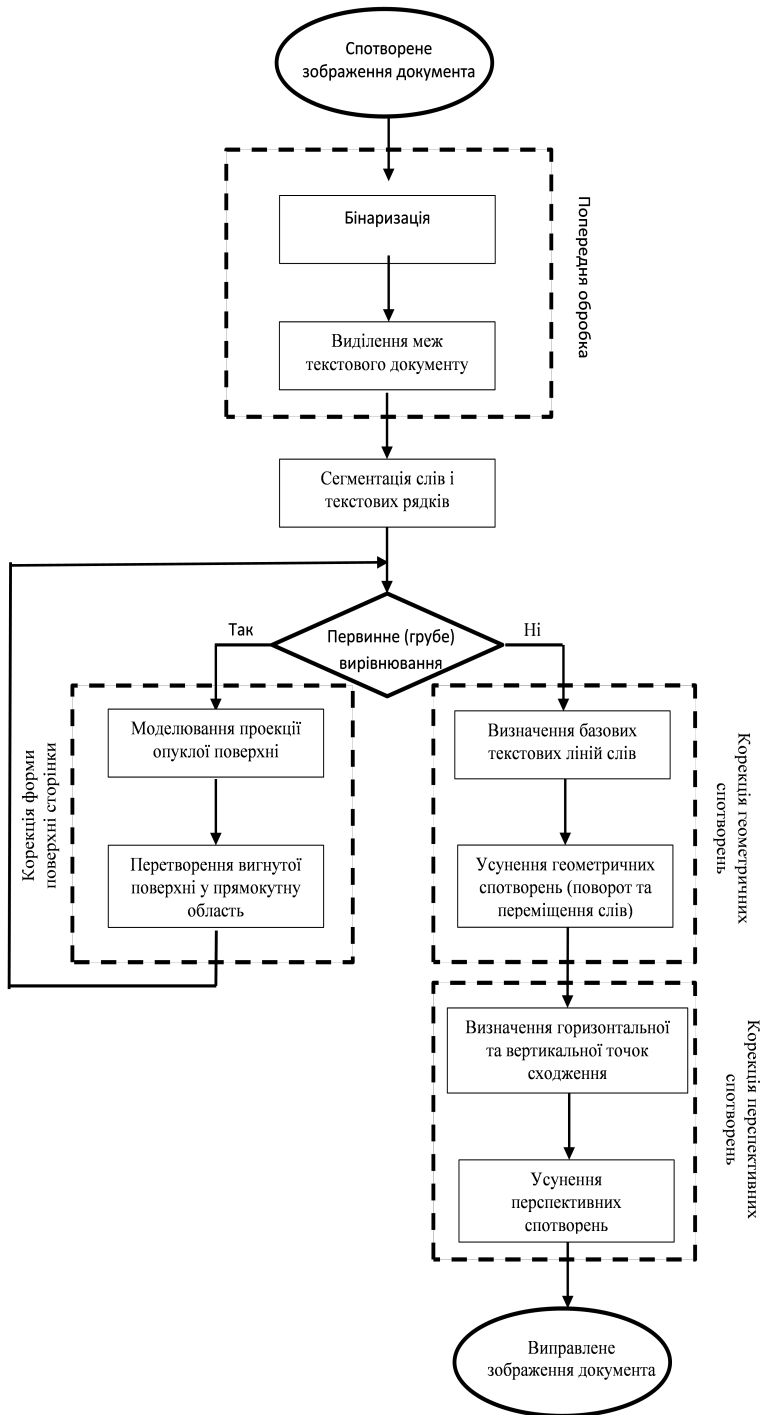


Рис.1. Загальна схема розробленого методу відновлення спотворених зображень текстових документів

Оцінка якості розпізнавання

На практиці неможливо отримати вихідний ідеальний неспотворений документ, а тому не можна точно оцінити ступінь розпрямлення зображення текстового документа. Відбувається це через те, що точно виміряти криві, які утворюють лінії тексту на зображенні, неможливо. Інакше кажучи, оскільки зображення документа завжди скановане не ідеально (на будь-якому растровому зображенні завжди присутні помилки пов'язані з дискретизацією), то не можна однозначно визначити ступінь прямої рядків на зображенні. Було розроблено спеціальні методи для оцінки якості корекції спотворених зображень [18, 19], оцінюючи рівень кривизни лінії рядка тексту на зображенні. Проте, як вже згадувалося, будь-яка оцінка прямої рядків, яку можна побудувати, буде неточною.

Кінцевою метою всіх методів корекції спотворень є покращення якості розпізнавання, а тому доцільно використати як оцінку системи OCR і порівнювати якість розпізнавання до і після виправлення спотворень зображень текстових документів.

Оцінкою точності розпізнавання можуть бути символи чи слова. Оцінка методів корекції спотворень зображень текстових документів з використанням OCR-систем виконується обчисленням точності розпізнавання символів чи слів у початковому (спотвореному) і виправленому документі після застосування одного з методів корекції спотворень. Для вимірювання точності використовуються символи і слова на оригінальних (спотворених) і виправлених зображеннях текстових документів [20].

Точність розпізнавання символів визначається за допомогою метрики (Word error gate) [21] і дорівнює відстані Левенштейна [22] між символами розпізнаного тексту і символами вихідного тексту. Тобто це мінімальна кількість операцій редагування (додавання символу (I), видалення з тексту символу (D) і заміни символу (S)), необхідна для приведення першого вигляду тексту до другого, поділена на кількість символів вихідного тексту (N):

$$WER = \frac{S + I + R}{N}. \quad (1)$$

Оцінка виконується шляхом розрахунку величини, яка визначає якість розпізнавання тексту до випрямлення рядків тексту, і величини, що характеризує якість розпізнавання тексту після випрямлення рядків. Оцінкою ефективності алгоритму випрямлення є величина, яка дорівнює різниці цих двох величин.

Експериментальні дослідження

Для оцінки якості розробленого методу було проведено серію експериментів. Тестовою базою для експерименту була публічна тестова база відсканованих зображень документів із геометричними спотвореннями [23] (25 зображень) та зображення, отримані авторами самостійно фотографуванням цифровою фотокамерою книг чи окремих аркушів текстових зображень (ще 25 зображень).

У публічну тестову базу [23] входять 102 зображення, з яких методом випадкового відбору для експерименту вибрали 25. Кожне зображення — це відскановане зображення однієї сторінки з будь-якої книги. Всі зображення чорно-білі.

На зображеннях текст розміщений в одну колонку і займає всю сторінку. Обсяг тексту на зображеннях у середньому становив 14 пунктів. Перші 11 зображень були відскановані з роздільною здатністю 300 точок на дюйм. Всі інші зображення відскановані з роздільною здатністю 450 точок на дюйм.

Решта 25 зображень — це зображення, сфотографовані цифровою фотокамерою. Процес отримання зображень виконувався відповідно до таких вимог:

- відстань між об'єктивом камери і поверхнею сторінки документа становить близько 30–60см;
- кут між об'єктивом камери і поверхнею книги становить не більше ніж 30 градусів;
- на зображеннях текст розміщений в одну колонку і займає всю сторінку;
- текст українською, російською чи англійською мовами.

Експеримент полягав ось у чому. Відбувається корекція спотворень кожного зображення з тестової бази за допомогою описаного алгоритму. Потім за допомогою програми OCR (вибрано ABBY FineReader 9.0 Professional) розпізнається вихідне зображення і зображення, отримане після корекції. Програма ABBY FineReader 9.0 була обрана за високу якість розпізнавання [24] (за різними оцінками це або найточніша система розпізнавання в світі для європейських мов, або друга за точністю після OmniPage). Точність обраної програми навіть на спотворених зображеннях тексту дає підстави сподіватися, що головним фактором, який впливає на якість розпізнавання документів з тестової бази, буде саме ступінь випрямлення рядків. Оцінювався відсоток виправлених рядків за рахунок корекції спотворень помилок розпізнавання.

Зображення текстових документів містять різні розміри шрифту, а також один чи декілька типів спотворень. Для порівняння до цих зображень було застосовано виправлення спотворень за допомогою комерційної програми BookRestorer.

На більшості зображень запропонований алгоритм істотно поліпшив якість розпізнавання — це видно з гістограми на рис. 2.

У середньому кількість помилок розпізнавання після застосування описаного алгоритму зменшується на 79 % (від кількості помилок на початковому викривленому зображенні). При цьому на вихідних зображеннях середній відсоток помилок розпізнавання дорівнював 43,60 %, а на розпрямлених за допомогою запропонованого методу зображеннях відсоток помилок розпізнавання становив 8,99 %. Зазначимо, що це істотне поліпшення якості розпізнавання спотворених зображень.

Отже, запропонований механізм корекції геометричних спотворень повністю довів свою ефективність на спотворених зображеннях текстових документів.

Крім того, у табл. 1 наведено загальні результати розпізнавання для всіх етапів розробленого методу. Як показали результати, застосування кожного етапу розробленого методу поліпшує якість розпізнавання зображень текстових документів і дає можливість досягти кращого результату, ніж за допомогою комерційної програми BookRestorer.

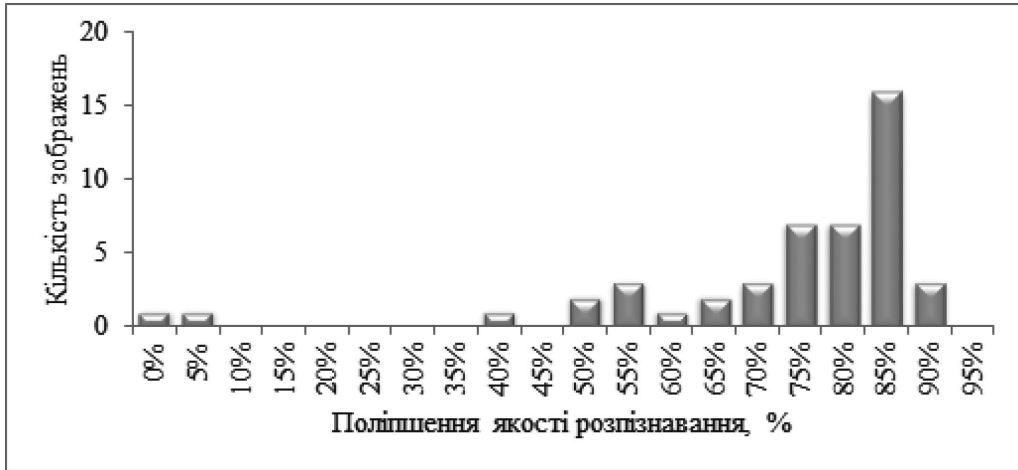


Рис. 2. Гістограма значень поліпшення якості розпізнавання зображень після застосування розробленого алгоритму корекції геометричних і перспективних спотворень

Таблиця 1

Середня точність розпізнавання, визначена із 50 спотворених зображень документів

Метод виправлення спотворень	Кількість символів	Кількість помилково розпізнаних символів	Точність розпізнавання
Без виправлення	90 272	39 358	56,40%
Після корекції форми поверхні сторінки шляхом побудови моделі перетворення поверхні сторінки	90 272	16 601	81,61%
Після корекції геометричних спотворень зображень тексту за допомогою виділення, повороту та переміщення зображень слів	90 272	12 855	85,76%
BookRestorer	90 272	15 260	83,10%
Розроблений метод	90 272	8 117	91,01%

На рис. 3 показано результат застосування всіх етапів методу. Запропонований метод коригує спотворення, здійснюючи вертикальне і горизонтальне вирівнювання документа. У прикладі показано, що можуть виникати деякі помилки в визначенні рядків тексту через неправильне визначення базової лінії слова. Проте ці помилки не впливають на результат процесу розпізнавання. Після корекції форми поверхні сторінки шляхом побудови моделі її перетворення (рис. 3б) виникає кілька помилок через неправильне визначення базових ліній слів, а також є багато спотворених слів

на правій стороні документа. Метод корекції геометричних спотворень зображень тексту за допомогою виділення, повороту та переміщення зображень (рис. 3в) не виправляє вертикальних спотворень зображень документів. Комерційна програма BookRestorer не може виправити всі наявні на зображенні деформації (рис. 3г), тому багато текстових рядків залишаються нерівними.

2.8. Character Cell Processing

Character cells are isolated within the spatial boundaries of each word. The paths of white space divide the word rectangle into preliminary character cells that extend from the top to the line bottom and are bounded on the left and right by inter-character spaces. These preliminary cells are later expanded to fill certain consecutive components. Note that in a small number of instances when the character is horizontally aligned, such as double quote or some Cyrillic or Greek characters, more than one character cell per character will be generated. In this case, the distance between these two cells, as opposed to the distance between adjacent cells, is usually detected. In Korean, the inter-component space is much smaller than the inter-character space.

4. Script Classification

Scripts are divided into two broad classes: Roman and Han-based, including Kana, Hingana, Katakana and Hangul. This gross classification is accomplished on the basis of the vertical distribution of upward concavities.

It is trivial to examine sets of runs within the connected component to determine the presence and location of upward concavities. Where two runs of black pixels appear on a single scan line of the raster image, if there is a run on the line below that spans the distance between these two runs, an upward concavity is formed on the line (Fig. 5). The reference coordinate system used for defining the spatial distribution of concavities is the character cell baseline.



Figure 6(a) shows the word Laboratory and the positions of the upward concavities. The a's show one upward concavity near the baseline, while the b's show two upward concavities near the baseline and one well above the baseline. Figure 6(b) shows the positions of upward concavities in a single Kana character. For a Han-based script, the occurrence of upward concavities is significantly different from that for Roman script [12]. Because the more complex characters incorporate more instances of enclosed white space, there are many more concavity occurrences per character and the spatial distribution of these concavities is more random, whereas for Roman scripts the distribution is bi-modal.

a)

2.8. Character Cell Processing

Character cells are isolated within the spatial boundaries of each word. The paths of white space divide the word rectangle into preliminary character cells that extend from the top to the line bottom and are bounded on the left and right by inter-character spaces. These preliminary cells are later expanded to fill certain consecutive components. Note that in a small number of instances when the character is horizontally aligned, such as double quote or some Cyrillic or Greek characters, more than one character cell per character will be generated. In this case, the distance between these two cells, as opposed to the distance between adjacent cells, is usually detected. In Korean, the inter-component space is much smaller than the inter-character space.

4. Script Classification

Scripts are divided into two broad classes: Roman and Han-based, including Kana, Hingana, Katakana and Hangul. This gross classification is accomplished on the basis of the vertical distribution of upward concavities.

It is trivial to examine sets of runs within the connected component to determine the presence and location of upward concavities. Where two runs of black pixels appear on a single scan line of the raster image, if there is a run on the line below that spans the distance between these two runs, an upward concavity is formed on the line (Fig. 5). The reference coordinate system used for defining the spatial distribution of concavities is the character cell baseline.

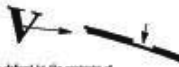


Figure 6(a) shows the word Laboratory and the positions of the upward concavities. The a's show two upward concavities near the baseline, while the b's show one concavity near the baseline and one well above the baseline. Figure 6(b) shows the positions of upward concavities in a single Kana character. For a Han-based script, the occurrence of upward concavities is significantly different from that for Roman script [12]. Because the more complex characters incorporate more instances of enclosed white space, there are many more concavity occurrences per character and the spatial distribution of these concavities is more random, whereas for Roman scripts the distribution is bi-modal.

b)

2.8. Character Cell Processing

Character cells are isolated within the spatial boundaries of each word. The paths of white space divide the word rectangle into preliminary character cells that extend from the top to the line bottom and are bounded on the left and right by inter-character spaces. These preliminary cells are later expanded to fill certain consecutive components. Note that in a small number of instances when the character is horizontally aligned, such as double quote or some Cyrillic or Greek characters, more than one character cell per character will be generated. In this case, the distance between these two cells, as opposed to the distance between adjacent cells, is usually detected. In Korean, the inter-component space is much smaller than the inter-character space.

4. Script Classification

Scripts are divided into two broad classes: Roman and Han-based, including Kana, Hingana, Katakana and Hangul. This gross classification is accomplished on the basis of the vertical distribution of upward concavities.

It is trivial to examine sets of runs within the connected component to determine the presence and location of upward concavities. Where two runs of black pixels appear on a single scan line of the raster image, if there is a run on the line below that spans the distance between these two runs, an upward concavity is formed on the line (Fig. 5). The reference coordinate system used for defining the spatial distribution of concavities is the character cell baseline.




Figure 6(a) shows the word Laboratory and the positions of the upward concavities. The a's show two upward concavities near the baseline, while the b's show one concavity near the baseline and one well above the baseline. Figure 6(b) shows the positions of upward concavities in a single Kana character. For a Han-based script, the occurrence of upward concavities is significantly different from that for Roman script [12]. Because the more complex characters incorporate more instances of enclosed white space, there are many more concavity occurrences per character and the spatial distribution of these concavities is more random, whereas for Roman scripts the distribution is bi-modal.

B)

2.8. Character Cell Processing

Character cells are isolated within the spatial boundaries of each word. The paths of white space divide the word rectangle into preliminary character cells that extend from the top to the line bottom and are bounded on the left and right by inter-character spaces. These preliminary cells are later expanded to fill certain consecutive components. Note that in a small number of instances when the character is horizontally aligned, such as double quote or some Cyrillic or Greek characters, more than one character cell per character will be generated. In this case, the distance between these two cells, as opposed to the distance between adjacent cells, is usually detected. In Korean, the inter-component space is much smaller than the inter-character space.

4. Script Classification

Scripts are divided into two broad classes: Roman and Han-based, including Kana, Hingana, Katakana and Hangul. This gross classification is accomplished on the basis of the vertical distribution of upward concavities.

It is trivial to examine sets of runs within the connected component to determine the presence and location of upward concavities. Where two runs of black pixels appear on a single scan line of the raster image, if there is a run on the line below that spans the distance between these two runs, an upward concavity is formed on the line (Fig. 5). The reference coordinate system used for defining the spatial distribution of concavities is the character cell baseline.




Figure 6(a) shows the word Laboratory and the positions of the upward concavities. The a's show two upward concavities near the baseline, while the b's show one concavity near the baseline and one well above the baseline. Figure 6(b) shows the positions of upward concavities in a single Kana character. For a Han-based script, the occurrence of upward concavities is significantly different from that for Roman script [12]. Because the more complex characters incorporate more instances of enclosed white space, there are many more concavity occurrences per character and the spatial distribution of these concavities is more random, whereas for Roman scripts the distribution is bi-modal.

Г)



д)

Рис. 3. Відновлення спотворених зображень документів: а) початкове зображення документа без виправлення; б) результат після корекції форми поверхні сторінки; в) після корекції геометричних спотворень зображень тексту за допомогою виділення, повороту та переміщення зображень слів; г) BookRestorer; д) запропонований метод

Висновки. Розроблений метод відновлення спотворених зображень документів працює з чорно-білими зображеннями, не потребує додаткового обладнання для відновлення форми поверхні сторінки, не залежить від типу спотворень і може бути застосований до зображень із комбінацією декількох типів спотворень. Процес сегментації текстових рядків проводиться на кожному етапі методу, і можемо зробити висновок, що точність сегментації на кожному етапі методу покращується (або в загальному не погіршується). Проведені експериментальні дослідження показали, що запропонований алгоритм забезпечує необхідну якість і обчислювальну ефективність. Доведено, що застосування розроблених методів корекції спотворень на етапі попередньої обробки перед розпізнаванням тексту дає змогу істотно підвищити якість розпізнавання. Рівень помилок зменшується з 43 % до 9 %, що є значним досягненням для систем цього класу. Зауважимо, що розроблений метод корекції геометричних та перспективних спотворень забезпечує якість попередньої обробки на вищому рівні, ніж комерційна програма BookRestorer.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Zhang L. An Improved Physically-Based Method for Geometric Restoration of Distorted Document Images // L. Zhang., Y. Zhang, C.L. Tan // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2008. — Vol. 30. — No. 4. — P. 728–734.
2. Ulges A. Document capture using stereo vision / A. Ulges, C. Lampert, T. M. Breuel // In Proceedings of the ACM Symposium on Document Engineering. — ACM. — 2004. — P. 198–200.

3. Yamashita A. Shape reconstruction and image restoration for non-flat surfaces of documents with a stereo vision system / A. Yamashita, A. Kawarago, T. Kaneko, K. T. Miura // Proceedings of 17th International Conference on Pattern Recognition (ICPR2004). — 2004. — Vol. 1. — P. 482–485.
4. Brown M. S. Document restoration using 3d shape: A general deskewing algorithm for arbitrarily warped documents / M. S. Brown, W. B. Seales // International Conference on Computer Vision (ICCV01). — 2001. — Vol. 2. — P. 367–374.
5. Cao H. Rectifying the bound document image captured by the camera: A model based approach / H. Cao, X. Ding, C. Liu // 7th International Conference on Document Analysis and Recognition. — Scotland. — 2003. — P. 71–75.
6. Liang J. Geometric rectification of camera-captured document images / J. Liang, D. DeMenthon, D. Doermann // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2008. — vol. 30. — no. 4. — P. 591–605.
7. Tan C. L. Restoring warped document images through 3D shape modeling / C. L. Tan, L. Zhang, Z. Zhang, T. Xia. — IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2006. — vol. 28. — no. 2. — pp. 195–208.
8. Zhang L. A unified framework for document restoration using inpainting and shape-from-shading // L. Zhang, A. M. Yip, M. S. Brown, C. L. Tan // Pattern Recognition. — 2009. — vol. 42. — no. 11. — P. 2961–2978.
9. 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 с.
10. Zhang Zh. Correcting Document Image Warping Based on Regression of Curved Text Lines / Zheng Zhang, Chew Lim Tan // Proc. ICDAR. — 2003. — P. 589–563.
11. Brown B. A model based book dewarping method using text line detection / Bin Brown, Minghui Wu, Rongfeng Li and other. — Proceedings of the Second International Workshop on Camera-Based Document Analysis and Recognition (CBDAR-2007). — Curitiba, Brazil, — 2007. — P. 63–70.
12. Mischke L. Document Image De-warping Based on Detection of Distorted Text Lines / L. Mischke, W. Luther // International Conference on Image Analysis and Processing. — Cagliari, Italy. — 2005. — P. 1068–1075.
13. Laviaille O. Active Contours Network to Straighten Distorted Text Lines // O. Laviaille, X. Molines, F. Angella, P. Baylou // International Conference on Image Processing. — Thessaloniki, Greece. — 2001. — P. 748–751.
14. Ulges A. Document image dewarping using robust estimation of curled text lines / A. Ulges, C. H. Lampert, T. M. Breuel. — 8th International Conference on Document Analysis and Recognition. — Seoul, Korea. — 2005. — P. 1001–1005.
15. Zhang Y. Arbitrary warped document image restoration based on segmentation and Thin-Plate Splines / Y. Zhang, C. Liu, X. Ding, Y. Zou. — 19 th International Conference on Pattern Recognition. — Florida, USA. — 2008. — P. 1–4.
16. Кульчицька І. О. Інформаційна технологія відновлення спотворених зображень текстових документів / І. О. Кульчицька, О. В. Тимченко, О. О. Тимченко // Моделювання та інформаційні технології. — 2015. — Вип. 75. — С. 69–79.

17. Кульчицька І. О. Особливості алгоритмів бінаризації зображень документів / І. О. Кульчицька, О. В. Тимченко // Зб. наук. пр. ІПМЕ НАН України. — Вип. 68. — К. : 2013. — С. 141–149.
18. Stamatopoulos B. A Methodology for Document Image Dewarping Techniques Performance Evaluation / B. Stamatopoulos, B. Gatos, I. Pratikakis // 10 th International Conference on Document Analysis and Recognition. — Barcelona, Spain. — 2009. — P. 956–960.
19. Fu B. A Document Rectification Approach Dealing With Both Perspective Distortion And Warping Based On Text Flow Curve Fitting / Bin Fu, Wenxin Li, Minghui Wu, Rongfeng Li // International Journal of Image and Graphics. — 2012. — Vol. 12. — No. 1. — P. 23–25.
20. Kanai J. Performance metrics for document understanding systems / J. Kanai, T. A. Nartker, S. Rice, G. Nagy // Proc. 2nd Int. Conf. Document Anal. Recognit. — Japan. — 1993. — P. 424–427.
21. Olive J. Handbook of Natural Language Processing and Machine Translation / Olive J. et al. // DARPA Global Autonomous Language Exploitation, Springer Science+ Business Media. — 2011.
22. Levenshtein V. I. Binary Codes Capable of Correcting Deletions, Insertions, and Reversals / V. I. Levenshtein // Soviet Physics Doklady, 10. — 1966. — P. 707–710.
23. Datasets University Of Kaiserslautern. — URL : <http://www.iupr.com/archived-2009/datasets>.
24. OCR Software Review. — URL : <http://ocr-software-review.toptenreviews.com>.

REFERENCES

1. Zhang, L., Zhang, Y., & Tan, C. L. (2008). An Improved Physically-Based Method for Geometric Restoration of Distorted Document Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 30, 4, 728–734 (in English).
2. Ulges, A., Lampert, C. & Breuel, T. M. (2004). Document capture using stereo vision. In *Proceedings of the ACM Symposium on Document Engineering*. ACM, 198–200 (in English).
3. Yamashita, A., Kawarago, A., Kaneko, T. & Miura, K.T. (2004). 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, 482–485 (in English).
4. Brown, M. S. & Seales, W. B. (2001). Document restoration using 3d shape: A general deskewing algorithm for arbitrarily warped documents. In *International Conference on Computer Vision (ICCV01)*, 2, 367–374 (in English).
5. H. Cao, X. Ding, & C. Liu. (2003). Rectifying the bound document image captured by the camera: A model based approach. *7th International Conference on Document Analysis and Recognition*, Scotland, pp. 71–75. (in English).
6. J. Liang, D. DeMenthon, & D. Doermann. (2008). Geometric rectification of camera-captured document images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, 4, pp. 591–605. (in English).
7. C.L. Tan, L. Zhang, Z. Zhang, & T. Xia. (2006). Restoring warped document images through 3D shape modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, 2, pp. 195–208. (in English).

8. L. Zhang, A.M. Yip, M.S. Brown and C.L. Tan. (2009). A unified framework for document restoration using inpainting and shape-from-shading. *Pattern Recognition*, vol. 42, 11, pp. 2961–2978. (in English).
9. A. Masalovich. (2007). Using Bezier patch to approximate the distortion of text documents. *Proceedings of the 17th International Conference on Computer Graphics and Vision (GraphiCon 2007)*, Moscow, pp. 239–243. (in English).
10. Zheng Zhang, Chew Lim Tan. (2003). Correcting Document Image Warping Based on Regression of Curved Text Lines. *Proc. ICDAR-2003*, pp.589–563. (in English).
11. Brown, Bin, Minghui, Wu, Rongfeng, Li, Wenxin, Li, Zhuoqun, Xu & Chunxu, Yang (2007). A model based book dewarping method using text line detection. *Proceedings of the Second International Workshop on Camera-Based Document Analysis and Recognition (CB-DAR-2007)*, 63–70 (in English).
12. Mischke, L. & Luther, W. (2005). Document Image De-warping Based on Detection of Distorted Text Lines. *International Conference on Image Analysis and Processing*, 1068–1075 (in English).
13. Laviaille, O., Molines, X., Angella, F. & Baylou, P. (2001). Active Contours Network to Straighten Distorted Text Lines. *International Conference on Image Processing*, 748–751 (in English).
14. Ulges, A., Lampert, C. H. & Breuel, T. M. (2005). Document image dewarping using robust estimation of curled text lines. *8th International Conference on Document Analysis and Recognition*, 1001–1005 (in English).
15. Zhang, Y., Liu, C., Ding, X. & Zou, Y. (2008). Arbitrary warped document image restoration based on segmentation and Thin-Plate Splines. *19th International Conference on Pattern Recognition*, 1–4 (in English).
16. Kulchytska, I. O., Tymchenko, O. V. & Tymchenko, O. O. (2015). *Informatsiina tekhnolohiia vidnovlennia spotvorenykh zobrazen tekstovykh dokumentiv. Modelling and Information Technologies*, 75, 69–79 (in Ukrainian).
17. Kulchytska, I. O., & Tymchenko, O. V. (2013). *Osoblyvosti alhorytmiv binaryzatsii zobrazen dokumentiv. Proceedings of IPM NAS of Ukraine*, 68, 141–149 (in Ukrainian).
18. Stamatopoulos, Gatos, B. & Pratikakis, I. (2009). A Methodology for Document Image Dewarping Techniques Performance Evaluation. *10th International Conference on Document Analysis and Recognition*, 956–960 (in English).
19. Bin, Fu, Wenxin, Li, Minghui, Wu & Rongfeng, Li (2012). Document Rectification Approach Dealing With Both Perspective Distortion And Warping Based On Text Flow Curve Fitting. *International Journal of Image and Graphics*, Vol. 12, 1, 23–25 (in English).
20. Kanai, J., Nartker, T. A., Rice, S., & Nagy, G. (1993). Performance metrics for document understanding systems. In *Proc. 2nd Int. Conf. Document Anal*, 424–427 (in English).
21. Olive, J. (2011). *Handbook of Natural Language Processing and Machine Translation: DARPA Global Autonomous Language Exploitation*, Springer Science+ Business Media (in English).
22. Levenshtein, V. I. (1966). Binary Codes Capable of Correcting Deletions, Insertions, and Reversals. *Soviet Physics Doklady*, 10:707–710 (in English).
23. Datasets University Of Kaiserslautern. Retrieved from: <http://www.iupr.com/archived-2009/datasets> (in English).
24. OCR Software Review. Retrieved from: <http://ocr-software-review.toptenreviews.com> (in English).

RESEARCH AND EVALUATION OF CORRECTION METHODS OF GEOMETRIC DISTORTIONS OF TEXT DOCUMENT IMAGES

O. V. Tymchenko ^{1,2}, I. O. Kulchytska ¹, R. O. Kulchytsky ¹, O. O. Tymchenko ¹

¹*Ukrainian Academy of Printing,
19, Pid Holoskom St., Lviv, 79020, Ukraine
irynekylch@gmail.com*

²*Uniwersytet Warmińsko-Mazurski,
2, Michała Oczapowskiego, Olsztyn, 10-719, Polska,
o_tymch@ukr.net*

We have analyzed the existing methods of correction of geometric distortions of text document images and reviewed the weaknesses of each method. The new method of distortion correction has been developed, that does not depend on the type of distortion and can be used to images of a combination of several types of distortions. We have examined the features of algorithms evaluation for correcting distortions of images using OCR. The experimental research has shown that the application of the developed methods for correcting distortions during pre-treatment before the text recognition can significantly improve the recognition quality. The experimental research has shown that the methods of correction of geometric and perspective distortions provide higher quality of pre-treatment levels than commercial software BookRestorer.

Keywords: *distorted images, OCR system, text document image, recognition quality, distortion correction, image preprocessing.*

Стаття надійшла до редакції 07.06.2016.

Received 07.06.2016.