

Видалення фону з зображень сторінок книг та документів

Ганна Терещенко, Сергій Вовк

Факультет фізики, електроніки та комп'ютерних систем

Дніпропетровський національний університет ім. О. Гончара, м. Дніпропетровськ

anyateres@gmail.com, vovk_s_m@mail.ru

Abstract

An approach to eliminating of background on the images of book pages and documents is suggested. This approach is based on simple methods of nonlinear spatial filtering, provides high speed and efficiency, and can be used in the software of mobile devices for processing of text images.

1. Вступ

В даний час комп'ютерна техніка і сучасні мобільні пристрої дозволяють робити швидкі копії документів. Але у багатьох випадках такі копії не мають належної якості, важкі для розпізнавання, читання та сприйняття. Це обумовлено тим, що основні способи отримання зображень сторінок книг і документів - сканування і фотографування - мають ряд істотних недоліків. Недоліки сканування книг полягають у появі темної смуги і геометричних спотворень у місці розвороту книги; крім того, якість сканування часто залежить від ступеня притиску кришки сканера. Недоліки фотографування документів пов'язані з відблисками, що виникають на поверхні книги від фотоспалахів або від яскравого джерела світла, а також з нерівномірною освітленістю сторінок, геометричними і перспективними спотвореннями рядків тексту. В цілому для практичного застосування зазначених способів отримання зображень текстів найбільш суттєвою проблемою є нерівномірний фон на зображеннях, і тому існує необхідність у розробці простих підходів для його усунення. Це дозволить більш ефективно використовувати зображення тексту для його розпізнавання в OCR-системах, при формуванні файлів для електронних книг, а також дозволить поліпшити візуальну якість текстів і формул при їх безпосередньому читанні і сприйнятті людиною.

2. Постановка задачі дослідження

Традиційний підхід до вирішення задач обробки зображень заснований на використанні методів просторової фільтрації [1-3]. Стосовно обробки зображень текстів книг і документів, для усунення нерівномірності фону можуть бути використані різні типи нелінійних фільтрів, зокрема: контргармонійні усереднюючі фільтри CH_p (Contra-harmonic mean), L_p -усереднюючі фільтри і фільтри на порядкових статистиках (Order Statistic Filters). Це обумовлено тим, що дані фільтри дозволяють ефективно усувати "білі" і "чорні" викиди яскравості на зображеннях, а також забезпечують низькі вимоги до обчислювальних ресурсів системи обробки і мають високу алгоритмічну ефективність.

Контргармонійні і L_p -усереднюючі фільтри відносяться до класу нелінійних усереднюючих фільтрів [3]. У

дискретному випадку ці фільтри описуються співвідношенням:

$$g(m,n) = U^{-1} \left[\frac{\sum_i \sum_j a(i,j) U[f(m+i,n+j)]}{\sum_i \sum_j a(i,j)} \right]; (i,j) \in D_z, (1)$$

де $g(m,n)$ - результуюче значення; $a(i,j)$ - вагові коефіцієнти фільтра, $U[\dots]$ - однозначна аналітична функція, яка задає нелінійне перетворення і має зворотну функцію $U^{-1}[\dots]$, $f(i,j)$ - значення вихідного зображення. Співвідношення (1) визначає наступний алгоритм. Спочатку всі значення вихідного зображення, які потрапили в задану локальну область D_z піддають нелінійному перетворенню $U[\dots]$, потім отримані значення множать на відповідні вагові коефіцієнти $a(i,j)$, підсумовують, утворюючи одне значення, яке ділять на суму вагових коефіцієнтів, після чого до отриманого значення застосовують нелінійне перетворення $U^{-1}[\dots]$.

Якщо $U[f] = (f)^p$, де $p \in \mathbb{Q} \setminus (-1,0,1)$, то з (1) отримуємо L_p -усереднюючий фільтр, який на виході дає так зване L_p -середнє g_{L_p} всіх вибірок вхідного зображення [3]. Наприклад, на виході одновимірного L_p -усереднюючого фільтра при $a(i) = \text{const}$ будемо мати:

$$g_{L_p} = \left(\frac{1}{N} \sum_{i=1}^N |f(i)|^p \right)^{1/p}, (2)$$

де N - розмір апертури фільтра. Якщо вагові коефіцієнти $a(i,j)$ не вибирати рівними константі, то можна отримати інший клас нелінійних усереднень. Одне з нелінійних усереднень такого класу є контргармонійне середнє [3]:

$$g_{CH_p} = \frac{\sum_i \sum_j f^{(p+1)}(i,j)}{\sum_i \sum_j f^p(i,j)}, (3)$$

яке може бути інтерпретоване як арифметичне середнє з вагами $a(i,j) = f^p(i,j)$.

Фільтри на порядкових статистиках є дискретними аналогами квантильної фільтрації неперервних зображень [2]. Поняття порядкової статистики застосовується до дискретної вибірки даних кінцевої довжини. Обчислення j -ої порядкової статистики здійснюється за наступним алгоритмом. Спочатку вихідна кінцева вибірка групується в порядку збільшення своїх значень; отримана впорядкована (або ранжирувана) у зростаючому порядку вибірка називається вектором порядкових статистик. Після цього вибирається j -й елемент, який використовується в якості результату, що виробляється фільтром. Перевага фільтрів на порядкових статистиках полягає в їх дуже ефективній програмній реалізації застосовно до цифрових зображень.

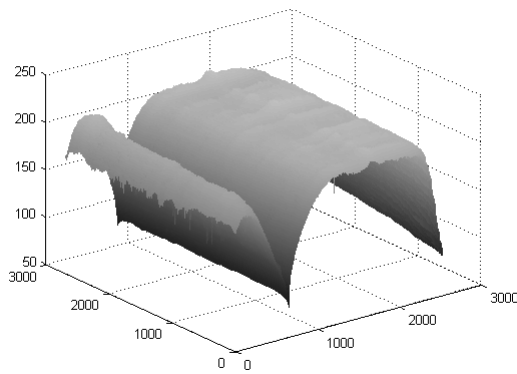
Задача досліджень полягала у визначенні можливостей та ефективності застосування зазначених нелінійних просторових фільтрів для вирішення задачі

усунення фону на зображеннях сторінок книг і документів.

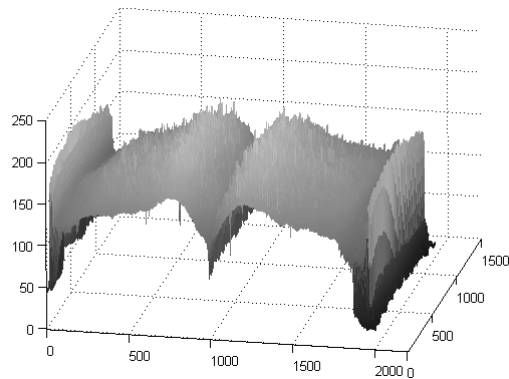
3. Рішення задачі і результати досліджень

Типові поверхні, які демонструють нерівномірність фону при скануванні і при фотографуванні книг, наведені на мал.1. Видно, що фон є вкрай нерівномірним вздовж рядків книги і приблизно постійним поперек рядків. Також чітко видно зміни фону в місцях згину книги і на краях книги. З розгляду двовимірних перерізів зображення книги вздовж і впоперек рядків (рис.2) також видно, що в напрямку поперек рядків можлива нескладна гладка

апроксимація (наприклад, сплайнами, відношенням поліномів і т. п.) поверхні книги (рис.1, а, б), тоді як гладка апроксимація вздовж рядків є проблематичною [4, 5]. З цієї причини має сенс виконувати обробку зображень книг поперек рядків, бо в цьому випадку для апроксимації буде потрібно невисокий ступінь апроксимуючих поліномів і, отже, обробка не буде надмірно трудомісткою. Оскільки зміну фону поперек рядків можна вважати приблизно рівною нулю в межах невеликих за розміром локальних областей, то нелінійна фільтрація може мати локальний характер, тобто може бути орієнтованою на отримання одного значення в межах кожної з цих локальних областей.



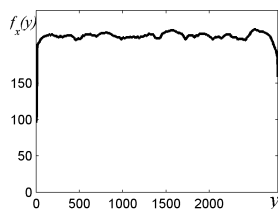
а)



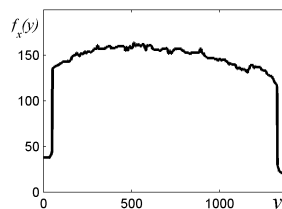
б)

Рис.1: Приклади тривимірних зображень поверхні книги:

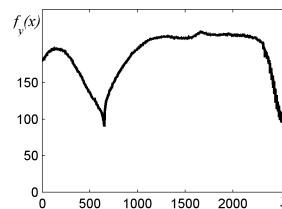
а) - при скануванні, б) - при фотографуванні.



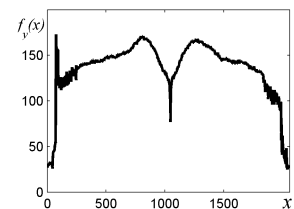
а)



б)



в)



г)

Рис.2: Двовимірні зображення перетину книги:

а), в) - поперек і вздовж рядків при скануванні; б), г) - поперек і вздовж рядків при фотографуванні,

В основу підходу, що пропонується, було покладено ідею використання одновимірного нелінійного просторового фільтра, орієнтованого в напрямку поперек рядків документа і такого, що ковзає по зображенню. У рамках даного підходу одним з важливих є питання про вибір типу фільтра та його довжини. Враховуючи особливості розв'язуваної задачі, можна запропонувати наступне рішення: довжина фільтра повинна бути достатньою для захоплення одного рядка тексту і одного міжрядкового інтервалу. Оскільки практично будь-яка книга чи документ містить на сторінці від 40 до 50 рядків, то довжину фільтра априорно можна вибрати приблизно рівною $L_y / 40$, де L_y - розмір зображення поперек рядків тексту.

Після застосування такого нелінійного фільтра символи тексту на отриманому зображенні будуть усунуті або значно розмиті; таке зображення далі буде трактуватися як зображення фону, що описує поверхню книги або документа та її певний окіл. Для отримання зображення тексту необхідно виконати віднімання зображення фону з вихідного зображення, а потім до кожного отриманого елемента зображення додати постійне значення, формуючи тим самим рівномірний фон бажаної інтенсивності (наприклад, можна вибрати середину динамічного діапазону яскравості).

Шляхом моделювання та обробки зображень з текстом були досліджені методи просторової фільтрації на основі контргармонійних і L_p -усереднюючих фільтрів для

значень p в діапазоні від 1 до 9, а також фільтрів на порядковій статистиці в діапазоні від 55 до 95 процентилів розподілу (рис. 3). В умовах помірного зашумлення тестового зображення найменше значення

середнього квадрату помилки було досягнуто для фільтра на порядковій статистиці, що відповідала значенням з діапазону від 70 до 75 процентилів розподілу.

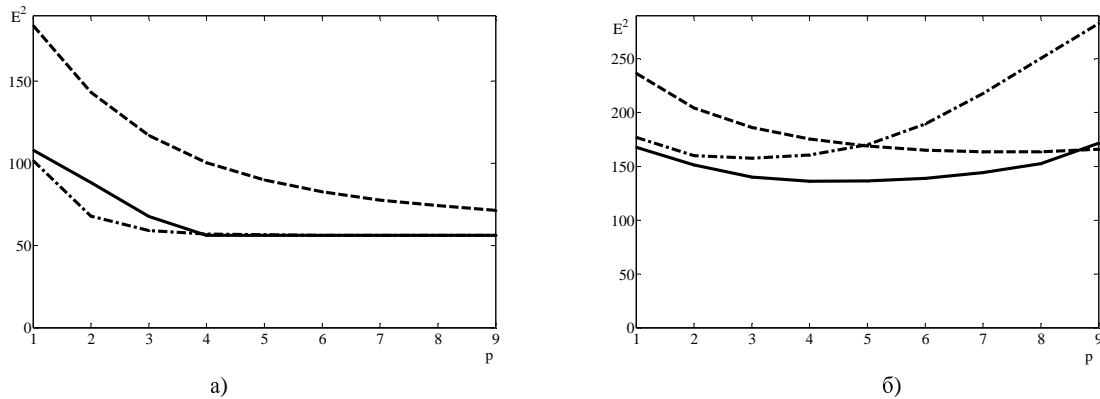


Рис 3. Залежність середнього квадрату помилки від параметрів фільтрів для випадків:
 а) тестове зображення без шуму; б) тестове зображення з адитивним та імпульсним шумом;
 суцільна крива – фільтр на порядковій статистиці, де $p=1$ відповідає 55 процентилю, а $p=9$ відповідає 95 процентилю; штрихова крива – L_p -усереднюючий фільтр; штрихпунктирна крива – контргармонійний фільтр.

Отриманий результат було підтверджено на практиці під час обробки реальних зображень сторінок просканованої книги та сфотографованої книги. Результати застосування фільтра на порядковій статистиці, що відповідала 75 процентилю розподілу, до зображення просканованої

книги наведені на рис. 4. Відзначимо, що недоліком цієї обробки є збільшення інтенсивності символів тексту в області згину сторінок, але цей недолік може бути усунутий шляхом додаткової обробки.

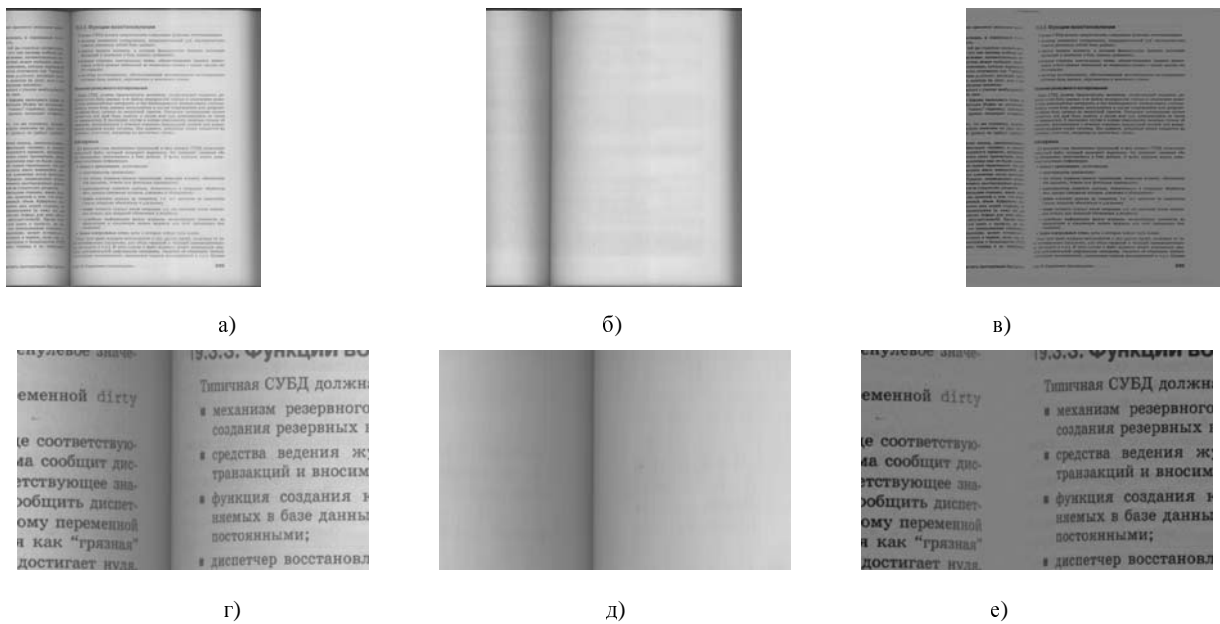
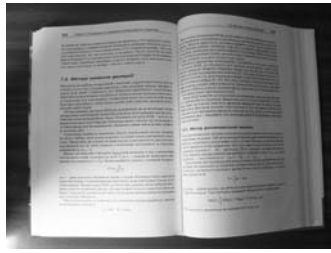


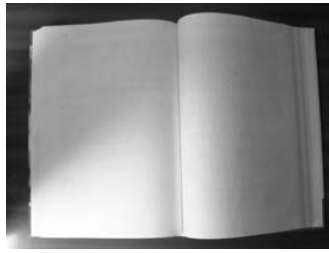
Рис.4: Видалення фону із зображення просканованої книги:
 а), г) - вихідне зображення і його фрагмент, б), д) - отриманий фон і його фрагмент, в), е) - результат усунення фону

Один з результатів обробки зображень сфотографованої книги наведений на рис. 5. Видно, що нерівномірний фон,

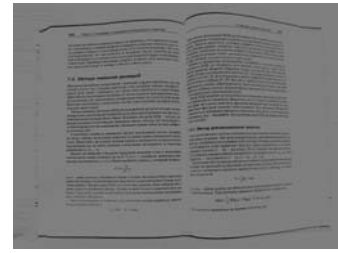
який був спричинений відблиском від освітлювального приладу, практично був усунутий.



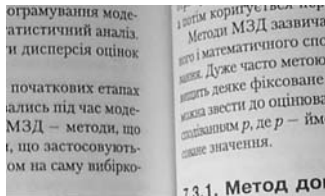
а)



б)



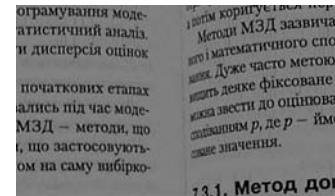
в)



г)



д)



е)

Рис. 5: Видалення фону із зображення сфотографованої книги:
а), г) - вихідне зображення і його фрагмент, б), д) - отриманий фон і його фрагмент,
в), е) - результат усунення фону

Ефективність роботи розроблених алгоритмів перевірялась часом обробки, який звичайно залежить як від складності алгоритму, так і від розміру зображення, що обробляється. Так, зображення, яке було отримане шляхом сканування (рис.4), мало розмір 2543x2789 пікселів, а час його обробки в середовищі MatLab на комп'ютері Celeron (R) CPU 2.0GHz склав 0,31 секунди. Зображення, яке було отримане шляхом фотографування (рис.5), мало розмір 2058x1379 пікселів, а час його обробки склав 0,15 секунди. Такі результати вказують на достатню ефективність підходу, що пропонується, та можливість його використання в мобільних пристроях з "повільними" процесорами. Зауважимо, що зазначений час обробки може бути суттєво скорочений за рахунок оптимізації програмного коду. Запропонований підхід був узагальнений на випадок кольорових зображень та полягав в їх поканальній обробці. Результати обробки кольорових зображень та їх аналіз представлені в доповіді.

4. Висновки

Для усунення з зображення нерівномірного фону, який описує поверхню книги при її скануванні або при її фотографуванні, досить ефективним є застосування одновимірних нелінійних просторових фільтрів, орієнтованих в напрямку поперек рядків тексту. Дослідження вказують на доцільність використання фільтру на порядковій статистиці, яка відповідає 75

процентилію розподілу в локальній околиці, що формується апертурою фільтра та захоплює один рядок тексту разом з міжрядковим інтервалом. Розроблений алгоритм може бути вбудований в програмне забезпечення мобільних пристроїв з "повільним" процесором для рішення задач обробки зображень сторінок книг та документів.

5. Література

- [1] Форсайт, Девид А., Понс, Жан. Компьютерное зрение. Современный подход.: Пер. с англ. – М.:Издательский дом «Вильямс», 2004. – 928 с.: ил. – Парал. тит. англ
- [2] Прэтт У. Цифровая обработка изображений. В 2 кн. - М.: Мир, 1982. Кн.1. - 312 с.; Кн.2, – 480 с.
- [3] Venetsanopoulos A.N. Digital image processing and applications / Department of Electrical and Computer Engineering, University of Toronto. – Canada, 1996. – 598 p
- [4] Зейналов Р. Ш., Геометрическая коррекция фотографий документов // Proc. Graphicon-2008, 2008, pp. 258-261
- [5] Ramiz Zeynalov, Alexander Velizhev, Anton Konushin. Document images geometrical distortions correction using text lines shape extraction, http://gc2009.graphicon.ru/files/gc2009/proceedings/conference/se9/2/2_Paper.pdf