

# АДАПТИВНЕ ПІДВИЩЕННЯ ЯСКРАВОСТІ ЗОБРАЖЕНЬ В ЗАДАЧАХ БІНАРИЗАЦІЇ ТЕКСТОВИХ ДОКУМЕНТІВ

**В.І. Загребнюк**

Кандидат технічних наук, доцент\*

Контактний тел.: (048) 723-12-84

E-mail: vampiter@rambler.ru

**Д.Д. Ленік**

Викладач\*

Контактний тел.: (048) 720-77-06

E-mail: ddlenik@ukr.net

Кафедра автоматизації та управління технологічними процесами  
Одеська національна академія зв'язку ім. О.С. Попова  
вул. Ковальська, 1, м. Одеса, Україна, 65020

*У цій статті наведені методи та алгоритми попередньої оцінки якості зображення та глобально-адаптивного підвищення яскравості. Показано, що зображення з підвищеною яскравістю можна бінаризувати з використанням простих глобальних порогових методів*

*Ключові слова: оцінка зображення, покращення, яскравість, адаптивна бінаризація, зображення документа*

*В этой статье представлены методы и алгоритмы предварительной оценки качества изображения и глобально-адаптивного повышения яркости. Показано, что изображения с повышенной яркостью можно бинаризовать с использованием простых глобальных пороговых методов*

*Ключевые слова: оценка изображения, улучшение, яркость, адаптивная бинаризация, изображение документа*

*Methods and algorithms of preliminary estimate of image quality and global adaptive increase of luminance are present in this paper. It is shown that image with the promoted luminance could be binarized with the use of simple global thresholding binarization methods*

*Key words: estimate image, enhancement, adaptive binarization, brightness, document image*

## 1. Введення

На сьогодні активно розвивається декілька нових напрямків стосовно розробки методів, алгоритмів та застосувань бінаризації, основні з них – розпізнавання текстів у мобільних терміналах, перетворення в електронну форму архівних рукописних документів, розпізнавання стародрукованих архівних документів тощо.

В залежності від характеристик вхідного зображення, задачу бінаризації можна розділити на наступні класи:

- бінаризація в задачах виділення текстових сегментів з кольоровими символами;
- бінаризація кольорових зображень документів з ахроматичним текстом;
- бінаризація текстових документів у градаціях сірого.

У кожному з цих класів задач розробляються свої методи та алгоритми. Найбільш широко бінаризація зображень використовується в системах оптичного розпізнавання символів (ОРС). Відомо, що надійність розпізнавання текстів суттєво залежить від якості

цифрового зображення документа. Стосовно ахроматичних зображень документів, якісним зображенням документу прийнято вважати таке, гістограма, або щільність розподілу рівнів сірого якого є бімодальною, тобто рівні сірого символів суттєво відрізняються від рівнів сірого фону, а сам фон є однорідним. На практиці якісне зображення скоріше всього є виключенням. У переважній більшості випадків зображення документів мають низьку якість: неоднорідний фон, наявність шумів, низький контраст символів тощо. Іншими словами, кожне зображення має свій, не бімодальний розподіл щільності градацій сірого або контекст. Все це надзвичайно ускладнює процедуру бінаризації.

Оскільки, не існує методів, що дозволяли б оцінити якість вхідного зображення, навіть, якщо воно представлено в градаціях сірого, то не зважаючи на те, що за двадцять років було запропоновано досить значну кількість методів та алгоритмів бінаризації текстових документів, на сьогодні, ця проблема не є остаточно вирішеною. Практично у всіх методах бінаризації, з наведених вище класів, використовуються від одного до декількох емпіричних параметрів, які підбираються в інтерактивному режимі, з урахуванням якості зо-

браження документу. Звичайно, що в такому випадку, задовільна якість бінаризації і подальше розпізнавання тексту досягається лише на обмеженій множині, близьких за контекстними характеристиками, зображень.

Звідси, стосовно задачі бінаризації можна виділити наступні актуальні проблеми, а саме:

- оцінка якості зображення документу та його покращення;
- розробка адаптивних методів бінаризації з урахуванням контексту зображення без використання емпіричних коефіцієнтів.

## 2. Аналіз існуючих методів бінаризації

В загальному методи бінаризації розділяють на дві групи: глобальна бінаризація та локальна бінаризація. У методах глобальної бінаризації значення порогу задається або розраховується для всього зображення. Глобальна бінаризація ефективна, коли гістограма вхідного зображення – бімодальна, тобто, коли яскравість символів достатньо чітко відрізняється від яскравості фону [1]. На практиці зображення з такими характеристиками можна отримати лише за умови високої якості самого документу.

У випадках, коли якість документу низька (рис. 1), на його зображенні рівні сірого змінюються в діапазоні  $I \in [I_{min}, I_{max}]$ , який менше динамічного діапазону рівнів сірого у обраній системі, присутні шуми, фон є неоднорідним, а контраст між символами та фоном - незначний. У таких випадках використовується переважно локально-адаптивна порогова бінаризація.

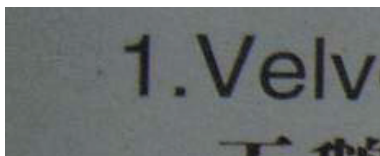


Рис. 1. Приклад зображення документу низької якості

Одними із перших методів локальної адаптивної порогової бінаризації є метод Ніблека [2]. У цьому методі значення порогу  $T$  обирається в межах певної околиці пікселу (маски) розміру  $N \times N$  ( $N = 2k + 1$ ), що бінаризується, виходячи зі співвідношення:

$$T = m + k \cdot \sigma$$

де  $m$  – середнє значення  $I$ , а  $\sigma$  – середньоквадратичне відхилення в межах маски;  $k$  – це емпірична константа, яка підбирається так, щоб забезпечити найбільш якісну бінаризацію. Даний метод в областях з плавною зміною рівнів сірого призводить до виникнення артефактів та шумів (рис. 2).

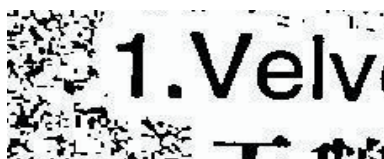


Рис. 2. Бінаризація з використанням методу Ніблека

Щоб усунути недоліки методу Ніблека, виходячи з припущення, що значення  $I$  символів тексту розташовані ближче до мінімально можливого значення (наприклад, до нуля), а значення  $I$  пікселів фону (не текстових областей) ближче до максимально можливого значення (наприклад, 255 в динамічному діапазоні  $[0, 255]$ ), у [3,4] запропоновано обчислювати локальне значення порогу з рівності:

$$T = m \cdot \left( 1 + k \cdot \left( \frac{\sigma}{R} - 1 \right) \right),$$

де  $R$  – середина динамічного діапазону рівнів сірого, а  $k$  – емпірична константа, значення якої підбирається для певного класу зображень з однаковим розподілом градацій сірого. Запропонований метод дає кращі результати у текстових областях, а що стосується фону, то як і у методі Ніблека при бінаризації виникають артефакти.

У [5] запропонована модифікація методу Саувола [3] з врахуванням локального, в межах маски, контрасту та максимального глобального контрасту. З урахуванням цього для визначення порогу використовується наступне співвідношення:

$$T = (1 - k) \cdot m + k \cdot I_{min}^{gl} + k \cdot \frac{\sigma}{\max(\sigma_i)} \cdot (m - I_{min}^{gl}),$$

де  $I_{min}^{gl}$  – глобальне мінімальне значення рівнів сірого;  $\max(\sigma_i)$  – максимальне значення середньоквадратичного відхилення по всім частинам зображення, що мають розмір маски;  $k$  – емпірична константа. Як і в розглянутих раніше методах, в області фону при невдалому виборі розміру маски, при бінаризації виникають артефакти та шуми.

У [6] запропонований локально адаптивний метод бінаризації, що враховує особливості розподілу градацій сірого в області рядків тексту. У даному методі бінаризація включає три етапи: початкова сегментація для виділення пікселів що належать рядкам тексту; покращення символів з використанням інформації про їх найближче оточення, контрастування та згладжування текстових областей; остаточна бінаризація.

У [7] розглянуто метод бінаризації з використанням кластерного аналізу. Запропонований метод, за ідейним змістом є алгоритмом К-середніх та використовується для кластеризації зображення у градаціях сірого на два кластери, за умови хорошої якості зображення. Метод полягає у ітераційному пошуку, з використанням гістограм, мінімуму функції:

$$f(m_1, m_2) = \sum_{i=1}^2 \sum_{l \in S_i} n_l (1 - m_i)^2,$$

де  $n_l$  – кількість пікселів, що мають  $l$ -й рівень градацій сірого  $l \in [0, L-1]$ ,  $l$  – кількість рівнів градацій сірого на зображенні;  $m_i$  – центр  $i$ -го кластера;  $S_i$  – кластери.

Ітерації виконуються до тих пір поки центри кластерів або функція  $f(m_1, m_2)$  не перестануть змінюватися. Для отриманої кластеризації визначається глобальне значення порогу

$$T = \frac{m_1^* + m_2^*}{2}.$$

де  $m_1^*$  та  $m_2^*$  – остаточні значення центри кластерів.

У [8] запропоновано метод бінаризації, що використовує процедуру попереднього покращення зображення. Ця процедура полягає в ітеративному підвищенні яскравості та еквалізації зображення. Підвищення яскравості здійснюється за відомою формулою:

$$I_{i+1}(x,y) = 1 - T_i + I_i(x,y),$$

тут  $T_i$  - середньоарифметичне значення нормованої яскравості всього зображення на  $i$ -й ітерації, або поріг для бінаризації, який динамічно змінюється;  $I_{i+1}(x,y)$ ,  $I_i(x,y)$  - нормовані значення яскравості з діапазону  $[0,1]$  на  $(i+1)$ -й та  $i$ -й ітераціях, відповідно. Після підвищення яскравості найменше значення рівнів сірого буде дорівнювати  $I_{i+1}^{\min} = 1 - T_i + I_{\min}$ , де  $I_{\min}$  - мінімальне значення до підвищення яскравості. Для того щоб забезпечити якомога більшу відмінність між фоном та текстом здійснюється еквалізація зображення за формулою:

$$\tilde{I}_{i+1}(x,y) = 1 - \frac{1 - I_{i+1}(x,y)}{1 - E_i},$$

де  $E_i$  – мінімально можливе значення яскравості.

Поріг для бінаризації динамічно перевизначається у ході покращення зображення. Як підкреслюють автори, неоложком методом є відсутність умови згідно з якою можна було б визначити, коли необхідно припинити ітераційне покращення зображення, тому для кожного класу зображень необхідно підбирати кількість ітерацій.

Наведені вище методи розроблені для бінаризації цифрових зображень документів, що характеризуються складним контекстом та мають низьку якість. Для забезпечення якісної бінаризації необхідно підбирати, для кожної сукупності близьких за контекстом зображень, таку комбінацію розміру маски та значень емпіричних констант, що забезпечує мінімізацію помилок при розпізнаванні тексту. Що стосується зображень документів хорошої якості, то для їх бінаризації достатньо використовувати прості методи бінаризації, наприклад, глобальну порогову бінаризацію, де у якості порогу обирається середнє значення рівнів сірого. Отже, більшість проблем бінаризації можна зняти, якщо попередньо встановлюється якість вхідного зображення документу. Якщо відома якість вхідного зображення, то по крайній мірі можна визначитись з вибором такого методу бінаризації, що в даних умовах є найбільш ефективним.

Тому метою даної роботи є розробка методу та алгоритму визначення якості зображення документу та її покращення, так щоб спростити подальшу бінаризацію.

### 3. Визначення якості зображення документу та її адаптивне підвищення

Для бінаризації цифрових зображень текстових документів, що у більшості випадків є ахроматичними, необхідно вирішити низку проблем. По перше: необхідно визначити, який з методів бінаризації слід

обрати, для того щоб забезпечити безпомилкове розпізнавання тексту.

У високоякісному зображенні документу, тобто зображенні, яке можна безпосередньо бінаризувати, яскравість однорідного фону повинна дорівнювати  $I_{\max}^{\text{cs}}$  - максимально можливою рівню сірого (яскравості) у обраній кольоровій системі. Будемо вважати, що символи тексту мають найменший рівень сірого  $I_{\min}^{\text{cs}}$ , або чорний колір –  $I_{\min}^{\text{cs}} = 0$ . На границях символів є пікселі у яких яскравість змінюється в деякому діапазоні так, щоб забезпечити згладжування границь та якісне візуальне сприйняття форми символу. Частина таких пікселів у кожному символі значно менше частки чорних пікселів. З урахуванням цього середнє значення яскравості  $M_{\text{hq}}$  буде пропорційним відношенню  $(n \cdot m - p) / (n \cdot m)$ , де  $n \cdot m$  - кількість пікселів на зображенні;  $p$  - кількість чорних пікселів. Тоді для оцінки середньої яскравості можна використовувати співвідношення:

$$M_{\text{hq}} \cong \frac{n \cdot m - p}{n \cdot m} I_{\max}^{\text{cs}}.$$

Знайдемо оцінку дисперсії:

$$\sigma^2 = \frac{1}{n \cdot m} \sum_{x=0}^{n-1} \sum_{y=0}^{m-1} (I_{xy} - M_{\text{hq}})^2 \cong \frac{p(n \cdot m - p)}{(n \cdot m)^2} (I_{\max}^{\text{cs}})^2.$$

Якщо врахувати той факт, що практично для всіх якісних зображень текстових документів  $p / (n \cdot m) \ll 1$ , то тоді  $M_{\text{hq}} \approx I_{\max}^{\text{cs}}$ , а

$$\sigma^2 \approx \frac{p}{n \cdot m} (I_{\max}^{\text{cs}})^2. \tag{1}$$

Звідси слідує, що для якісних ахроматичних зображень документів  $M_{\text{hq}}$  є практично інваріантною, а дисперсія пропорційна щільності чорних пікселів.

Виходячи з отриманих результатів сформулюємо умову, при виконанні якої якість зображення є достатньою для безпосередньої бінаризації

$$M_d \in \left[ I_{\max}^{\text{cs}}, I_{\max}^{\text{cs}} \left( 1 - \sqrt{\frac{p}{n \cdot m}} \right) \right], \tag{2}$$

де  $M_d$  - середня яскравість зображення документу.

У тому випадку, коли мінімальне значення  $I_{\min}^{\text{cs}}$  яскравості символів тексту  $I_{\min}^{\text{cs}} > I_{\min}^{\text{c}}$ , то необхідно виконати перетворення:

$$\tilde{I}_{xy} = \frac{I_{\max}^{\text{cs}}}{I_{\max}^{\text{cs}} - I_{\min}^{\text{c}}} (I_{xy} - I_{\min}^{\text{c}})$$

та для визначення якості зображення скористатися формулою (2).

Якщо для деякого зображення  $M_d$  не задовольняє умові (2), то доцільно виконати процедуру покращення якості зображення. У якості такої процедури розглянемо ітераційне підвищення яскравості та еквалізацію зображення.

Для підвищення яскравості зображення на першому етапі обчислюється середнє значення яскравості  $M$  та встановлюється її максимальне значення  $I_{\max}$ . Збільшимо глобально яскравість зображення використовуючи наступне лінійне перетворення:

$$I_{xy}^{k+1} = I_{\max}^k - M_k + I_{xy}^k$$

де  $I_{xy}^k$  – яскравість пікселу  $(x,y)$  на  $k$ -й ітерації,  $I_{xy}^{k+1}$  – яскравість цього ж піксела на  $k+1$ -й ітерації. Після такого перетворення, середнє значення яскравості на  $k+1$ -й ітерації  $M_{k+1} = I_{\max}^k$ , а яскравість пікселів зображення буде змінюватись у діапазоні  $[I_{\min}^{k+1}, I_{\max}^{k+1}]$ , де  $I_{\min}^{k+1} = I_{\max}^k + I_{\min}^k - M_k$ . Що стосується  $I_{\max}^{k+1}$ , то її значення буде визначатись зі співвідношення:

$$I_{\max}^{k+1} = \begin{cases} 2I_{\max}^k - M_k, & 2I_{\max}^k - M_k < I_{\max}^{cs} \\ I_{\max}^{cs}, & 2I_{\max}^k - M_k \geq I_{\max}^{cs} \end{cases}$$

Тепер необхідно виконати перетворення еквалізації гістограми згідно співвідношення:

$$\tilde{I}_{xy}^{k+1} = \frac{I_{\max}^{k+1}}{I_{\max}^{k+1} - I_{\min}^{k+1}} (I_{xy}^{k+1} - I_{\min}^{k+1})$$

Процедура покращення зображення закінчується, коли дисперсія (1) перестане змінюватись, тобто буде виконуватись наступна умова:

$$P_k = P_{k+1}$$

Сформулюємо алгоритм визначення якості документу та її адаптивного покращення. Він буде складатись з наступних кроків.

Крок 1. Обчислити мінімальне  $I_{\min}^d$  та максимальне  $I_{\max}^d$  значення яскравості зображення документу та задати початкове значення лічильника ітерацій  $k=1$ .

Крок 2. Якщо виконується одна з умов  $I_{\min}^d > I_{\min}^{cs}$  або  $I_{\max}^d < I_{\max}^{cs}$ , то виконати лінійне перетворення

$$\tilde{I}_{xy} = \frac{I_{\max}^{cs}}{I_{\max}^d - I_{\min}^d} (I_{xy} - I_{\min}^d)$$

Крок 3. Обчислити середнє значення  $M_d$  та підрахувати кількість чорних пікселів  $p_k$ .

Крок 4. Якщо  $M_d > I_{\max}^{cs} \left(1 - \sqrt{\frac{p}{n \cdot m}}\right)$ , то алгоритм закінчено: зображення можна бінаризувати.

Крок 5. Запам'ятаємо значення  $P = p_k$ .

Крок 6. Підвищити яскравість зображення згідно формули  $I_{xy}^{k+1} = I_{\max}^k - M_d + I_{xy}^k$ , та знайти  $I_{\max}^{k+1}, I_{\min}^{k+1}$ .

Крок 7. Виконати перетворення еквалізації

$$\tilde{I}_{xy}^{k+1} = \frac{I_{\max}^{k+1}}{I_{\max}^{k+1} - I_{\min}^{k+1}} (I_{xy}^{k+1} - I_{\min}^{k+1})$$

Крок 8. Обчислити середнє значення  $M_d$  та кількість чорних пікселів на перетвореному зображенні  $P_{k+1}$ .

Крок 9. Якщо  $P_{k+1} \neq P$ , то  $P = P_{k+1}$ ,  $k = k+1$  та перейти до кроку 6, інакше алгоритм закінчено. Зображення можна бінаризувати.

З використанням цього алгоритму було розроблене застосування щодо покращення якості зображення та його глобальної бінаризації з порогом, що дорівнює  $M_d$ .

На рис. 3 наведені результати бінаризації різних за контекстом зображень

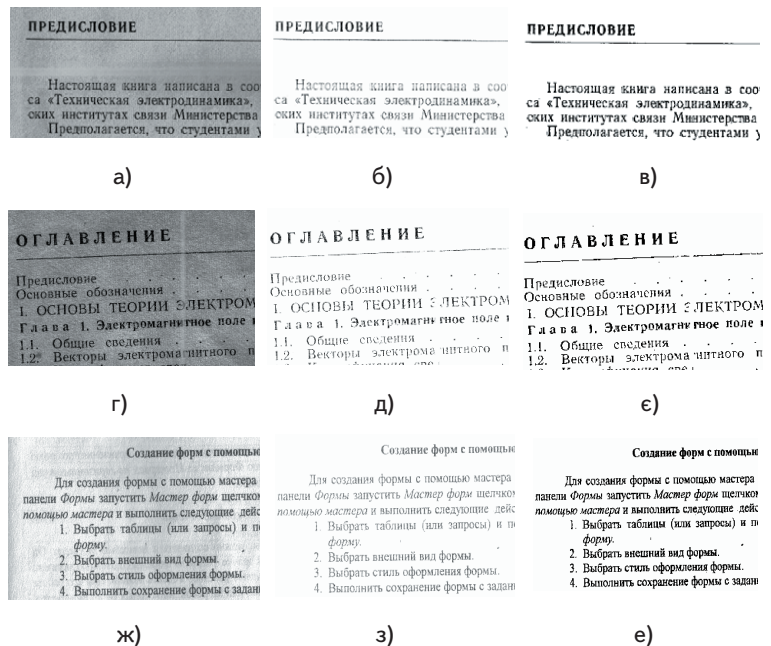


Рис. 3. Результати покращення та бінаризації зображень: а), б), в) – вхідні зображення; г), д), е) – покращені, в градаціях сірого; ж), з), е) – бінаризовані

Наведені зображення (рис. 3) розрізняються за контекстом, а саме: у зображення в першому рядку середнє значення рівнів сірого  $M_{d1} = 148,59$ , середньоквадратичне відхилення  $\sigma_{d1} = 35,17$  максимальна яскравість  $I_{\max1} = 230$ , а у зображення з другого рядка  $M_{d2} = 128,63$ ,  $\sigma_{d2} = 32,65$  та  $I_{\max2} = 218$ . Обидва зображення мають однакову мінімальну яскравість. У зображенні третього рядка –  $M_{d3} = 205,83$ ,  $\sigma_{d3} = 43,63$  та  $I_{\max3} = 255$ ,  $I_{\min3} = 6$ .

Для перевірки ефективності методу було оброблено 100 зображень документів отриманих зі сканеру з розподільними здатностями 300, 600 та 1200 dpi.

У всіх випадках, коли різні частини зображення практично не розрізняються за розподілом рівнів сірого, були отримані задовільні результати покращення якості зображення та його бінаризації.

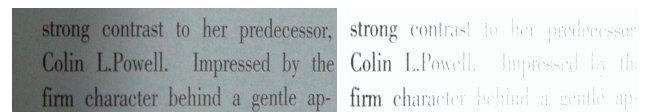


Рис. 4. Результати покращення зображення

На рис. 4 наведено зображення у якому можна виділити дві частини з різним контекстом. При глобальному покращенні якості такого зображення текст у його правій частині втрачається, оскільки символи мають більші значення рівнів сірого ніж символи у лівій частині зображення.

На рис. 5 наведені результати бінаризації правої частини зображення з рис. 4, а на рис. 6 – його лівої частини.

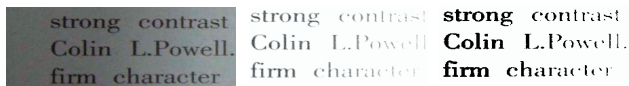


Рис. 5. Результати покращення та бінаризації лівої частини зображення рис. 4: зліва – вхідне зображення; посередині – покращене, в градаціях сірого; справа – бінаризоване

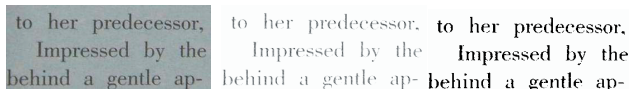


Рис. 6. Результати покращення та бінаризації правої частини зображення рис. 4: зліва – вхідне зображення; посередині – покращене, в градаціях сірого; справа – бінаризоване

Права частина зображення має такі характеристики:  $M_d = 107,87$ ,  $\sigma_d = 23,73$ ,  $I_{max} = 165$  та  $I_{min} = 15$ , а ліва частина зображення (див. рис. 4) –  $M_d = 158,9$ ,  $\sigma_d = 22,73$ ,  $I_{max} = 202$  та  $I_{min} = 44$ .

Якщо обробляти ці частини зображення окремо, то отримуємо задовільні результати покращення їх

якості та бінаризації. Тому подальше удосконалення запропонованого методу повинно полягати у розробці методів більш детального аналізу контексту зображення та локально-адаптивного підвищення яскравості.

### Висновок

За результатами виконаних досліджень можна зробити наступні висновки:

– запропонований метод попередньої оцінки якості зображення дозволяє визначити зображення які мають практично бімодальну гістограму рівнів сірого та можуть бути бінаризовані з використанням глобального порогового методу;

– результатом адаптивного підвищення яскравості зображень низької якості, частини якого не суттєво розрізняються за контекстом, є зображення з однорідним фоном та бімодальною гістограмою рівнів сірого;

– у тому випадку, коли різні частини зображення суттєво розрізняються за контекстом, адаптивне підвищення яскравості може супроводжуватись втратами корисної інформації.

### Література

1. Trier D. Evaluation of Binarization Methods for Document Images / D. Trier, T. Taxt // IEEE Trans. Pattern Analysis and Machine Intelligence. – 1995. – vol. 17, pp. 312-315.
2. Niblack W. An introduction to digital image processing / W. Niblack // Englewood Cliffs Prentice Hall, N.J. – 1986. – pp. 115-116.
3. Sauvola J. Adaptive document image binarization / J. Sauvola, M. Pietikainen // Patt. Recogn. – 2000 – vol. 33, pp. 225-236.
4. Zhang Z. Restoration of images scanned from thick bound documents / Z. Zhang and C. L. Tan // Proc. Int. conf. Image Processing. – 2001. – vol. 1, pp.1074-1077.
5. Wolf C. Extraction and recognition of artificial text in multimedia documents. / C. Wolf and J.-M. Jolion // Patt. Anal. Appl. – 2004. – vol. 6, no. 4, pp. 309-326.
6. Leedham G. Comparison of Some Thresholding Algorithms for Text/Background Segmentation in Difficult Document Images. / Graham Leedham, Chen Yan, Kalyan Takru, Joie Hadi Nata Tan and Li Mian // School of Computer Engineering – 639798.
7. Feng M. Contrast adaptive binarization of low quality document images / Meng-Ling Feng, Yap-Peng Tan // IEICE Electronics Express – 2004. – Val.1, No. 16. – 501-506.
8. Zhu Y. Document Image Binarization Based on Stroke Enhancement / Yuanping Zhu, Chunheng Wang, Ruwei Dai // Proceedings of the 18th International Conference on Pattern Recognition. – 2006.
9. Dong L. An Optimization-Based Approach to Image Binarization / Liju Dong, Ge Yu // Proceedings of the Fourth International Conference on Computer and Information Technology. – 2004.
10. Kavallieratou E. A Binarization Algorithm specialized on Document Images and Photos / Ergina Kavallieratou // Proceedings of the 2005 Eight International Conference on Document Analysis and Recognition. – 2005.