

П. М. Єгоров, О. І. Яковченко

Науково-дослідний, проектно-конструкторський
та технологічний інститут мікрографії (НДІ мікрографії)
пров. Академіка Підгорного, 1/60, 61046 Харків, Україна

Спосіб адаптивної бінарizaції напівтонових зображень з додатковим обробленням на основі відношення сигнал/шум

Запропоновано спосіб бінарizaції напівтонових зображень оцифрованих документів, який ґрунтується на використанні найбільш інформативних рівнів вейвлет-перетворення. Остаточне оброблення виконується з використанням класифікатора за відношенням сигнал/шум. Розроблено відповідне програмне забезпечення. Тестування показало, що запропонований спосіб забезпечує стабільні результати бінарizaції для доволі широкого за якісними ознаками класу зображень документів, включаючи зображення зі слабким контрастом і суттєвими ознаками згасання.

Ключові слова: бінарizaція зображень документів, вейвлет-перетворення, кореляційне зіставлення зображень, відношення сигнал/шум.

Вступ

На сьогодні у світі значної актуальності набрав переклад паперових документів у цифрову форму. До оцифрування надаються матеріали найрізноманітніших областей: музейні та архівні фонди, технічна документація промислових підприємств, що створена в паперовому вигляді, офісні документи під час створення баз цифрового документообігу та ще багато інших видів документів у різних сферах діяльності людини. Наприклад, один із найбільш амбіційних проєктів сучасності — оцифрування бібліотеки Ватикану [1].

Природно, що оцифровані документи, перед тим як стати доступними користувачам, підлягають тому чи іншому ступеню оброблення. Однією із найбільш поширених задач на цьому етапі є бінарizaція зображень у напівтоновому поданні. Бінарizaція у простішому випадку задається такою характеристичною функцією $\mu_{x,y}$:

$$\mu_{x,y} = \begin{cases} 1, & \text{якщо } I_{x,y} > \tau, \\ 0, & \text{якщо } I_{x,y} \leq \tau. \end{cases} \quad (1)$$

Суть у такому: якщо яскравість I пікселя з координатами (x, y) у вихідному зображенні перевищує деякий визначений поріг τ , то цей піксель належить до фону та приймає значення одиниці в бінаризованому зображенні. У протилежному випадку піксель належить до інформаційно значущої частини вихідного зображення, а в бінаризованому зображенні приймає значення нуль.

У більш загальному випадку пікселю надається значення характерної властивості (ознаки) зображення, наприклад деякої порівняльної статистичної характеристики яскравості в його околі на вихідному зображенні. Але, незалежно від способу визначення ознаки, залишається незмінне одне — наявність певного порогу такого значення і його визначення є найбільш важливою проблемою у бінарizaції напівтонових зображень.

Особливістю використання бінарizaції є те, що в багатьох випадках таке перетворення є ключовою частиною обробки. Так, бінарizaція має самостійне значення як кінцевий результат оброблення, наприклад, під час виокремлення певних об'єктів на зображенні [2, 3]. Не менш важливим є використання бінарizaції як проміжного етапу обробки, наприклад, у системах комп'ютерного зору [4] або під час проведення OCR (Optical Character Recognition — оптичного розпізнавання символів) [5].

Незважаючи на велику кількість запропонованих способів [6–9], на сьогодні універсальні алгоритми бінарizaції не знайдені. Такий стан є результатом різноманіття особливостей різних за походженням видів документів і цілей використання їхніх цифрових зображень. Відмітимо, практика свідчить про те, що конкретні методи є найкращими тільки для певних видів зображень. Тому увагу у цій статті зосереджено на важливому класі документів — оцифрованих текстових документах і конструкторській документації (далі — зображення документів). Актуальність бінарizaції текстової документації підтверджує проведення з цього приводу престижного міжнародного конкурсу DIBCO (Competition on Document Image Binarization) [10].

Сутність запропонованого способу бінарizaції

У даній роботі запропоновано спосіб бінарizaції, що спирається на виділення ознак масштабу за допомогою багатомасштабного аналізу, а саме вейвлет-аналізу. Вихідні положення такі:

— зображення документа є поєднанням трьох складових: інформаційної, шумової і фонові (головні складові якості);

— до інформаційної складової зображення в цій роботі віднесено області зображення, які містять графічні об'єкти, значущі для розпізнавання людиною змісту документа (далі — символи);

— до фонові складової належать всі області зображення, які не містять символів. У представленні, яке використовується в роботі, в області локалізації фонові складової значення яскравості не змінюються або змінюються значно повільніше відносно змін яскравості в області локалізації символів, тому ними можна знехтувати;

— до шумові складової віднесено випадкові відхилення значень яскравості від ідеалізованої незмінної фонові складової.

Головна ідея способу бінаризації така: за правильного вибору масштабу інформативних рівнів вейвлет-перетворення (тобто за ознаками масштабу) та використання порогових ознак яскравості можна виокремити область, де локалізована інформаційна складова зображення.

Вейвлет-аналіз — відносно новий, але вже вельми широко вживаний математичний апарат з достатньо повним освітленням у науковій літературі, тому не має сенсу зупинятися на його основах. Але необхідно відмітити деякі його важливі для цієї роботи особливості. Достатньо повне уявлення про нього наведено в огляді [11]. У зазначеному огляді його властивості представлено, виходячи з аналітичного подання. Комп'ютерне обчислення спирається на матричне подання. Відзначимо, що матричне подання має більшу повноту в тому сенсі, що містить об'єкти, які не мають аналітичної форми. У матричному поданні одновимірні вейвлети представлені парою векторів апроксимуючої і деталізуючої компонент (відповідно φ і ψ). Для вейвлет-перетворення двовимірних масивів застосовуються двовимірні вейвлети. У матричному поданні двовимірні вейвлети теж є матрицями. Матриця двовимірного вейвлета складається з чотирьох блоків і має такий вигляд:

$$\begin{bmatrix} \varphi \times \varphi & \psi \times \varphi \\ \varphi \times \psi & \psi \times \psi \end{bmatrix}. \quad (2)$$

У кожному блоці матриці (2) знаходиться компонента (теж матриця) двовимірного вейвлета розміром $M \times M$, утворена з двох компонент одновимірного вейвлета φ та ψ розміром M . У вейвлет-аналізі прийнята така інтерпретація: $\varphi \times \varphi$ — апроксимуюча компонента; $\psi \times \varphi$ — вертикальна деталізуюча компонента; $\varphi \times \psi$ — горизонтальна деталізуюча компонента; $\psi \times \psi$ — діагональна деталізуюча компонента. Вейвлет-перетворення функції, що аналізують, виконується окремо кожною компонентою. Вейвлет-перетворення має ієрархічний характер: рівень розкладення зростає відповідно до збільшення масштабу. Перший рівень відповідає найменшому масштабу. Вейвлет-перетворення є оборотним, тобто має пряме (декомпозиція) та зворотне (реконструкція) перетворення.

Однією з головних задач, яку необхідно вирішити, є визначення найбільш інформативних рівнів вейвлет-перетворення. Для цього використано математичний апарат, який розроблено щодо важливої області використання вейвлетів — стиснення зображень, і має назву вейвлет-пакетів [12]. У [13] звернено увагу на те, що в науковій літературі цей напрямок висвітлено недостатньо, тому наведемо деякі пояснення.

З точки зору теорії сигналів будь-який вейвлет є смуговим фільтром. За певних умов смуговий фільтр можна представити у вигляді пари низькочастотного (далі — НЧ) та високочастотного (далі — ВЧ) фільтрів. Важливою властивістю вейвлетів є те, що багато з них допускають таке подання.

Основним видом фільтрів, найбільш часто використовуваних у вейвлет-аналізі, є так звані квадратурно-дзеркальні фільтри (далі — КДФ). Опис процесу аналізу фільтрації сигналів за використанням КДФ наводять у вигляді дерева.

На рис. 1 наведено типовий приклад повного дерева вейвлет-перетворення (тобто до деякого обраного найбільшим рівня).

Гілки дерева відображають перетворення відповідно до визначеної компоненти: напрямком фільтрації компонентою низьких частот, напрямком фільтрації компонентою високих частот. Вузли відповідають рівню вейвлет-перетворення.

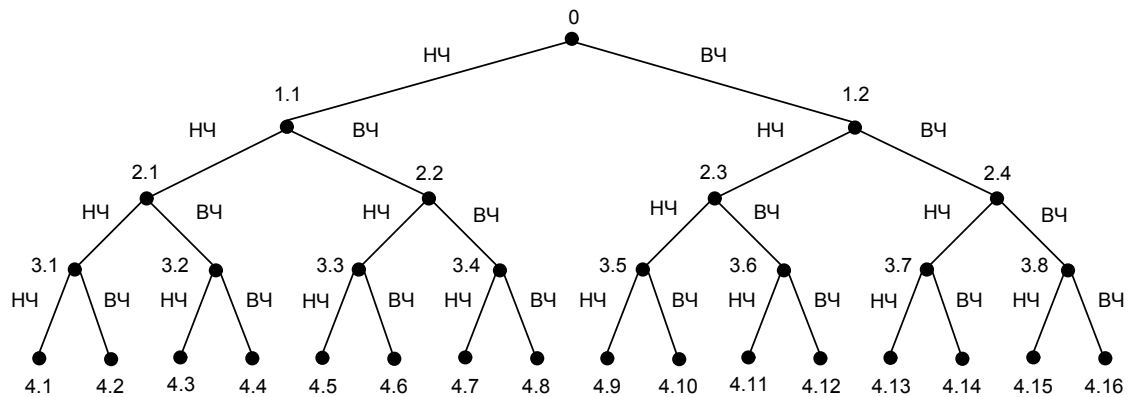


Рис. 1. Приклад дерева повного вейвлет-перетворення

Задачу можливо визначити як побудову найкращого дерева (або рівня) відповідно до визначеного критерію.

Оцінювання за обраним критерієм повинно відрізнити інформативність попереднього та наступного рівнів перетворення. Тобто, якщо на дереві уздовж гілки проміж двома вузлами змінення інформативності нижче за деякий поріг, то гілку до менш інформативного вузла буде відрізано (дерево не подовжується). За використання цього методу можна визначити вигляд найкращого дерева. Приклад можливого вигляду найкращого дерева наведено на рис. 2.

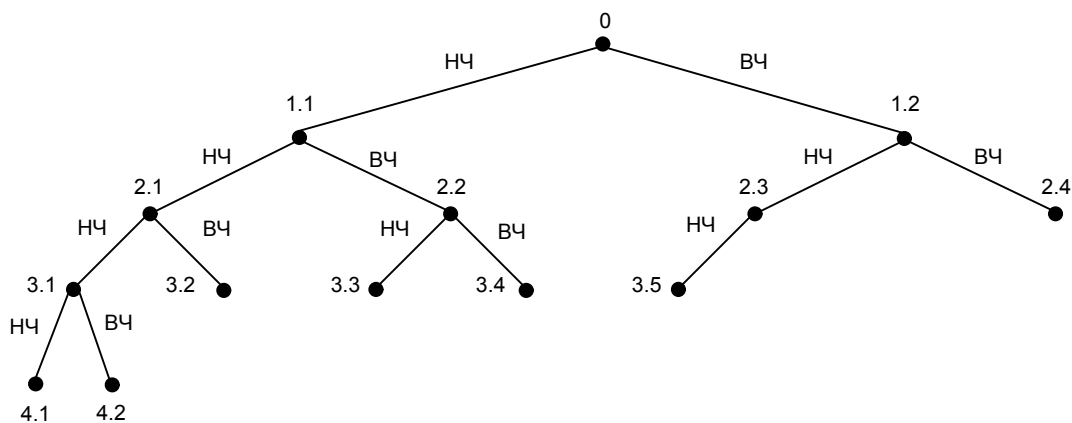


Рис. 2. Приклад можливого вигляду найкращого дерева вейвлет-перетворення

Найкращий рівень розкладення визначається за аналогією, але оцінювання здійснюється з використанням підсумованих показників порівнюваних рівнів у цілому. Найкраще за рівнем дерево, наприклад, порівняно з наведеним на рис. 1 повним, може містити три рівні перетворення.

Відмітимо, що особливістю відбору найбільш інформативних рівнів для реконструкції є те, що використовуються тільки два найбільші рівні, тобто менші видаляються як неінформативні. Також не використовується (видаляється) апроксимуюча складова вейвлет-образу.

У цій роботі використано вид ентропії, специфічний для вейвлет-аналізу — логарифм енергії (log energy) E_{\log} [12], яку обчислюють за формулою

$$E_{\log} = \log_2 \left(W_{\Delta,k}^2 \right), \quad (3)$$

де $W_{\Delta,k}^2$ — середнє значення квадратів вейвлет-коефіцієнтів по області Δ ; k — рівень вейвлет-перетворення.

Це поняття доцільно визначити більш докладно. Трактуювання наведено згідно з [14].

Використання статистичних методів у цифровій обробці зображень, зазвичай, обґрунтовується тим, що зображення представлено як випадковий процес. Із виникненням фрактального аналізу такий підхід набув уточнення: зображення можна представити як фрактальний випадковий процес.

Поняття фрактального випадкового процесу можливо пояснити на такому прикладі. Головною ознакою фрактального процесу є збереження самоподібності за зміною масштабу аналізу досліджуваного об'єкта. Припустимо, що аналізованим є деяке зображення. Із цього зображення обирають декілька однакових за розміром вибірок значень яскравості пікселей. Вибірки різняться в тому, що середні відстані між обраними пікселями відрізняються за значенням від одної вибірки до другої, наприклад, відповідно до співвідношення 1:2:4. Якщо між вибірками є значуща кореляція, то можливо стверджувати наявність фрактальних ознак.

Ступінь фрактальності випадкових процесів часто оцінюється показником Херста H . Значення цього показника належать до інтервалу між 0,5 і 1. Якщо $H = 0,5$, то це звичайний випадковий процес, а якщо більше, то фрактальний. Наприклад, багато природних процесів мають значення показника Херста близьке до 0,7.

Існує зв'язок між вейвлет-перетворенням і показником Херста:

$$\log_2 \left(W_{\Delta,k}^2 \right) \sim k^{(2H+1)}. \quad (4)$$

Тобто, ентропію логарифм енергії можливо трактувати як фрактальну ознаку зображення.

Метод визначення найкращого рівня вейвлет-перетворення здійснюється згідно з таким правилом: якщо ентропія логарифм енергії наступного рівня розкладення не перевищує поточний, то поточний рівень визначається як найкращий. У роботі [15] докладно доведено, що використання цієї ознаки до оброблення рентгенографічних зображень надає результат, найбільш відповідний до зорового сприйняття якості зображень.

Необхідно відмітити, що вейвлет-пакети використано тільки під час визначення найбільш інформативних рівнів декомпозиції. Одна з причин у тому, що існуючі алгоритми обчислення вейвлет-пакетної декомпозиції використовують дискретне вейвлет-перетворення (тут і далі — термінологія за пакетом Matlab),

тобто коли крок вікна зміщення в програмній реалізації точно дорівнює масштабу розкладення. За цієї умови під час реконструкції окремих рівнів у зображенні виникають артефакти, пов'язані із так званим «стробоскопічним ефектом» (видалення синхронних із сіткою вікон зміщення просторових частот у зображенні), що в цьому випадку є неприйнятним. Тому було використано тільки безперервне вейвлет-перетворення, тобто коли крок вікна зміщення у програмній реалізації дорівнює одному пікселю. Також було використано вейвлети Хаара та Добеші 2-го порядку, як найбільш прості з точки зору їхнього обчислення (вейвлети Хаара та Добеші 1-го порядку є тотожними). Однак досвід використання вейвлетів Хаара показав, що в реконструкціях, які містять обмежену кількість рівнів розкладення, доволі часто виникає фрагментація зображення на області квадратної форми різних розмірів. Це є наслідком характерної ступінчастої форми вейвлетів Хаара. З точки зору бінарizaції зображення за порогових ознак це може призвести до суттєвих помилок у відображенні форми об'єктів. Тому перевагу було віддано вейвлетам Добеші 2-го порядку, які мають більш дифузну форму (більш гладкі) [16] і помітної фрагментації не викликають.

На рис. 3 наведено загальний вигляд зображення конструкторського кресленника та його фрагмент, який використовується як приклад для оброблення.

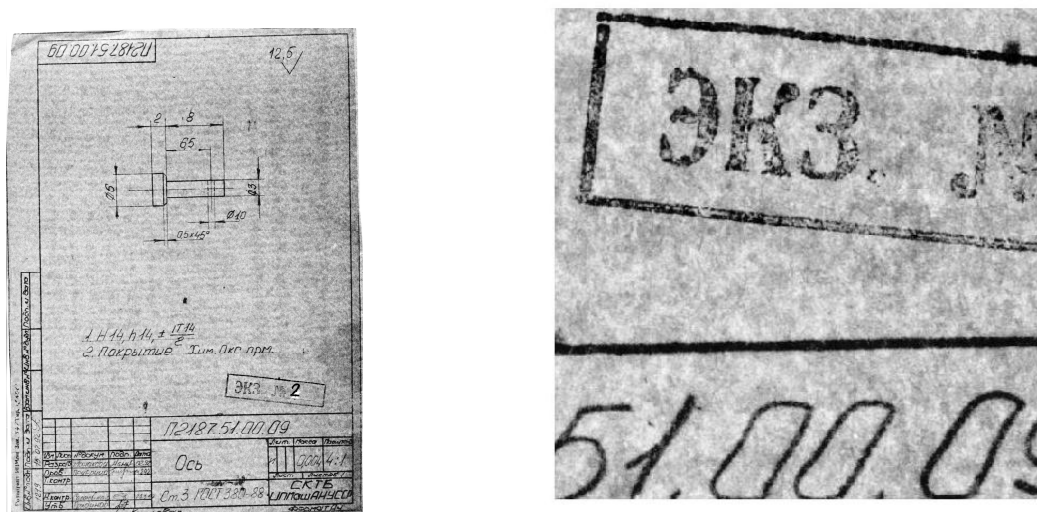


Рис. 3. Загальний вигляд зображення конструкторського кресленника та його фрагмент

Процес бінарizaції має ітеративний характер і складається з трьох кроків наближення.

На першому кроці наближення виконується первинна бінарizaція зображення:

- визначення найбільш інформативних рівнів декомпозиції;
- вейвлет-декомпозиція зображення за визначеними рівнями;
- сумісна реконструкція найбільш інформативних рівнів декомпозиції.

Відмітимо, що смугова фільтрація призводить до видалення постійної складової. Це суттєво у цьому випадку — видаляються повільні зміни яскравості фонові складові, пов'язані з нерівномірністю освітлення документа на зображенні.

Повільні в тому сенсі, що характерний масштаб змін значно більший від масштабу вейвлета. Тому у реконструкції середнє значення фоновї складовї дорівнює нулю.

Для первинної бінаризації використано такі належності до інформаційної складовї зображення:

— якщо значення матричного елемента (пікселя) з координатами (x, y) в реконструкції інформативних рівнів розкладення менше від значення «0», то відповідний піксель бінаризованого зображення (з координатами (x, y)) дорівнює нулю;

— якщо інакше, то відповідний піксель бінаризованого зображення дорівнює одиниці.

Тобто інформаційна складова (наприклад текст) буде чорною, а фонові складова — білою.

Наступний крок наближення — визначення верхнього порогу значень яскравості інформаційної складовї. Треба звернути увагу на таке: якщо значення порогу строго менше за нуль, то множина пікселів зображення, виділених на цьому кроці, містить як підмножину всю можливу до виділення за обраним способом інформаційну частину пікселів, а також підмножину хибного виділення пікселів фоновї складовї, як інформаційної складовї зображення.

Хибне виділення здебільшого пов'язане явищем, яке має аналогію в теорії сигналів — перехідними процесами під час фільтрації імпульсних сигналів. Суть у тому, що виникають коливання значень матричних елементів у реконструкції поблизу меж символів. Амплітуда коливань має як позитивні, так і негативні значення, тому в областях негативних значень бінаризація буде хибною — фоновї складовї надано значення «0». Фонові складова має яскравість більшу за інформаційну, тому хибне виділення можливо видалити за ознаками яскравості. Для цього необхідно провести порівняння інформаційних областей, виділених за першим кроком, з відповідними областями вихідного зображення. Якщо серед них присутні пікселі, яскравість яких перевищує деяку певну величину (тут — верхній поріг значень яскравості інформаційної складовї), то їм надається значення одиниці. В результаті відбувається уточнення належності до фоновї складовї бінаризованого зображення.

Визначення верхнього порогу значень яскравості інформаційної складовї виконано в такий спосіб: виконується ітеративний підбір значення порогу, за якого бінаризоване зображення з уточненою належністю до фоновї складовї здобуває найбільшої подібності до вихідного.

Метрикою подібності зображень обрано кореляцію Cr двох зображень, яка обчислюється за формулою

$$Cr = \frac{\sum_x \sum_y (I_{x,y}^{(A)} - I_{cp}^{(A)}) \times (I_{x,y}^{(B)} - I_{cp}^{(B)})}{\sqrt{\sum_x \sum_y (I_{x,y}^{(A)} - I_{cp}^{(A)})^2 \times \sum_x \sum_y (I_{x,y}^{(B)} - I_{cp}^{(B)})^2}}, \quad (5)$$

де x, y — координати пікселів зображення;

$I_{x,y}^{(A)}$ — значення яскравості пікселя зображення A з координатами (x, y) ;

$I_{cp}^{(A)}$ — середнє значення яскравості зображення A ;

$I_{x,y}^{(B)}$ – значення яскравості пікселя зображення B з координатами (x, y) ;

$I_{cp}^{(B)}$ — середнє значення яскравості зображення B .

Вибір такої метрики зумовлено позитивним досвідом її використання в технології DIC (Digital Image Correlation — цифрова кореляція зображень) [17], а також проведеними дослідженнями. На рис. 4 наведено значення кореляції Cr бінаризованого зображення другого кроку наближення та вихідного зображення залежно від значення яскравості I , яка обирається за верхній поріг значень яскравості інформаційної складової. У графіку є наявний максимум значень кореляції.

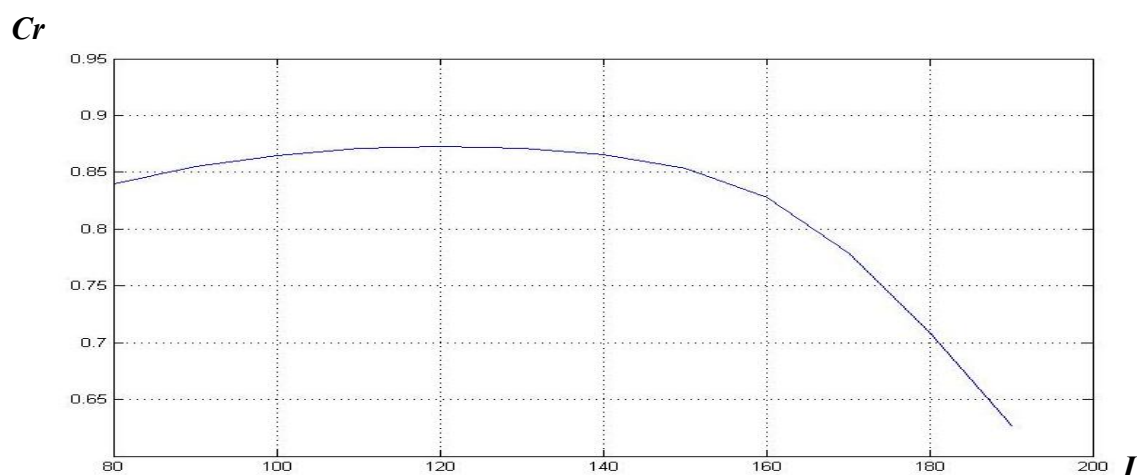


Рис. 4. Значення кореляції бінаризованого зображення інформаційної складової другого кроку наближення і вихідного зображення залежно від верхнього порогу значень яскравості інформаційної складової

На рис. 5 наведено результат бінарizaції фрагменту конструкторського кресленника.

На третьому кроці наближення виконується ще одне уточнення належності до фонові складової. У документах є графічні об'єкти зі значними характерними розмірами, наприклад стовпчикові діаграми. Всередині такі об'єкти мають характерне значення яскравості (наприклад середнє) менше, ніж характерні значення яскравості фонові складової. У межах такого об'єкта значення яскравості значно не змінюються. За останнього їхню внутрішню частину було класифіковано ще на першому кроці наближення як неінформаційну частину зображення (надано значення одиниці). Тому визначення

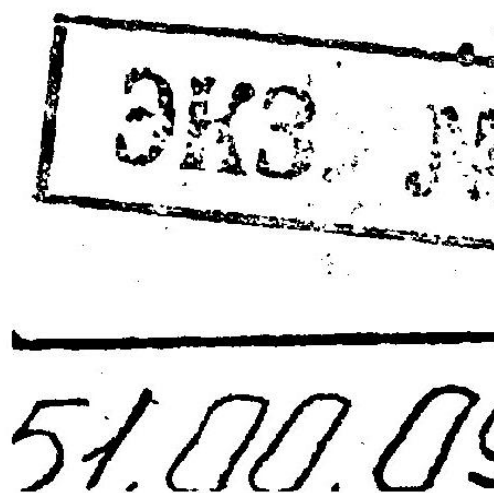


Рис. 5. Бінаризоване зображення інформаційної складової, в якому поріг другого кроку наближення визначено за результатами використання кореляційної метрики

в одержаній неінформаційній частині зображення об'єктів зі значеннями яскравості, меншими від верхнього порогу значень яскравості інформаційної складової, і є вирішенням зазначеної проблеми.

Далі наведено приклад застосування правила третього кроку наближення. Для виділення інформаційної складової використовується реконструкція тільки найменшого рівня вейвлет-перетворення. Відзначимо, що цей рівень виключено з реконструкції як неінформативний. Але в цьому випадку важливіше те, що за його використанням можливо виділити тільки контурні ознаки аналізованого зображення.

На рис. 6 ліворуч наведено результат бінаризації за другим кроком наближення, а праворуч — його уточнення за третім кроком наближення. Недоліком одержаного бінаризованого зображення є те, що в ньому більшою мірою визначається морфологічна складова шуму порівняно з використанням реконструкції найбільш інформативних рівнів вейвлет-перетворення.

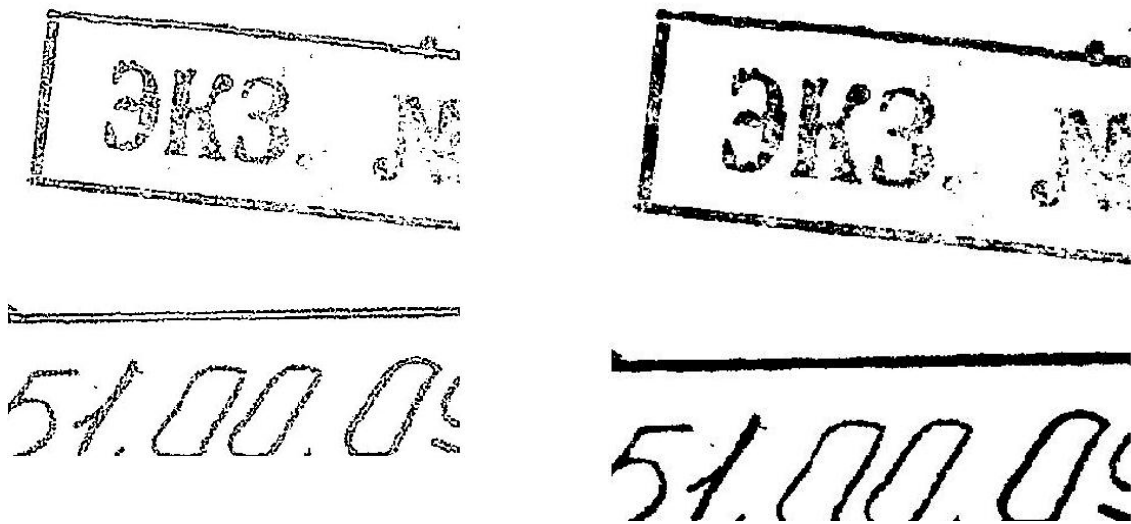


Рис. 6. Бінаризоване зображення інформаційної складової другого кроку наближення, в якому характерні розміри символів значно перевищують масштаб використовуваних вейвлетів

Таким чином, виділення інформативної складової зображення здійснюється з використанням трьох кроків наближення:

— значення порогу першого кроку наближення визначається за ознакою строго менше за нуль у значеннях найбільш інформативних рівнів вейвлет-перетворення;

— значення порогу другого кроку наближення визначається з використанням максимуму ознаки кореляції вихідного зображення та результату першого кроку наближення;

— для об'єктів значних розмірів визначення здійснюється з виділенням у неінформаційній (перші два кроки наближення) частині об'єктів зі значеннями яскравості, меншими за верхній поріг яскравості інформативної складової.

Додаткове оброблення

Запропонований алгоритм бінарizaції має особливість — якщо у площині локальної області, що надана до аналізу, відсутні символи, то наведений вище алгоритм бінаризує фонову текстуру зображення документу (див. рисунки далі), що не є прийнятним. Тому виникає необхідність проведення додаткового оброблення, або як зветься в літературі таке перетворення — «постпроцесінг» [7]. Для подолання вказаної проблеми використано додаткову класифікацію локальних областей. За основу класифікації взято локальне відношення сигнал/шум, яке позитивно зарекомендувало себе на практиці [18]. Відповідно до визначеного в цій роботі підходу в області локалізації фонові складові значення яскравості не змінюються або змінюються значно повільніше відносно змін яскравості в області локалізації символів. Тому природно метрикою значення F фонові складові обрати середнє значення яскравості пікселів по аналізованій області оброблення. Особливістю є те, що середнє обчислюється тільки для пікселів, які не належать до інформаційної складові:

$$F = \frac{\sum_x \sum_y (I_{x,y}^{(F)})}{Q_F}, \quad (6)$$

де $I_{x,y}^{(F)}$ — значення яскравості пікселя фонові складові зображення з координатами (x, y) ; Q_F — кількість пікселів фонові складові зображення.

Необхідно зазначити, що такий спосіб визначення фонові складові дає значно менше зміщення оцінки середнього порівняно з обчисленням оцінки середнього без видалення інформативної складові. За значного зміщення оцінки значення фону визначення метрик, які введені далі, втрачає сенс.

Сигнал визначено як різницю значення фонові складові та локального значення яскравості інформаційної складові зображення. Тоді просторовий розподіл сигналів визначено за формулою

$$S_{x,y} = F - I_{x,y}^{(V)}, \quad (7)$$

де $I_{x,y}^{(V)}$ — значення яскравості пікселя інформаційної складові зображення з координатами (x, y) .

До шумові складові віднесено випадкові відхилення значень яскравості пікселів, не належних до інформативної частини зображення, від значення фонові складові. Стандартною метрикою величини шуму є середньоквадратичне відхилення, тому значення шумові складові визначено за формулою:

$$N = \sqrt{\frac{\sum_x \sum_y (I_{x,y}^{(F)} - F)^2}{Q_F}}. \quad (8)$$

Метрику локального відношення сигнал/шум (тобто просторовий розподіл відношення сигнал/шум) $SN_{x,y}$ визначено як

$$SN_{x,y} = \log_2 \left(\frac{S_{x,y}}{N} \right). \quad (9)$$

У прикладних галузях науки та техніки логарифмічні метрики є стандартною формою визначення відношення сигнал/шум в інформаційних показниках. Відмітимо, це споріднює їх з інформаційною ентропією. Оскільки використовується двійкове логарифмування, то одиницю вимірювання метрик можливо трактувати як біт. Крім того, логарифмічне подання надає змоги більш детально відобразити області невеликих значень відношення сигнал/шум, що важливо для зображень низької контрастності.

Практичні результати

На основі наведеного способу розроблено відповідне програмне забезпечення (далі — ПЗ). Для проведення додаткового оброблення діапазон значень $SN_{x,y}$ розділено на п'ять інтервалів. За замовчуванням встановлено такі межі діапазонів (класів якості): 0–2, 2–3, 3–4, 4–5, 5–8 (верхнє значення належить до поточного діапазону). Інтервали якості відображаються в колірному поданні: білий, магента, зелений, синій, чорний відповідно. Межі інтервалів визначають у текстовому конфігураційному файлі. Таке подання надає можливості під час зорового оцінюванні колірному представлення швидко підібрати значення меж класифікації. Підбір необхідних значень виконується з використанням тестової вибірки зображень документів. У ПЗ мається можливість виключення та включення до бінаризованого зображення довільних окремих класів якості або їхні довільної комбінації.

Відмітимо таке: основний інтерес представляє бінаризація неякісних (слабких за контрастом) зображень документів, особливо згасаючих текстів. На рис. 7 наведено приклад бінаризації фрагменту сканованого зображення електронної схеми. Вихідний паперовий документ — електрографічна копія («білка») 1992 року з помітними ознаками згасання.

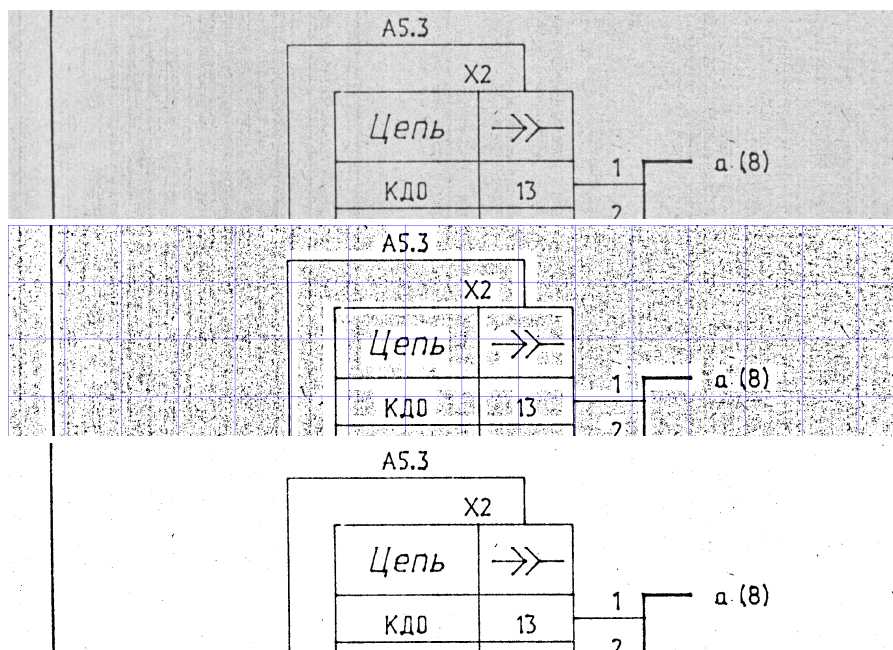


Рис. 7. Приклад бінаризації фрагменту сканованого зображення електронної схеми

Зверху — вихідне скановане зображення, у середині — бінаризоване без додаткового оброблення з бінаризованою фоновою текстурою, знизу — зображення з додатковим обробленням.

На рис. 8 наведено приклад бінарizaції конкурсного зображення документа з сайту DIBCO [10].

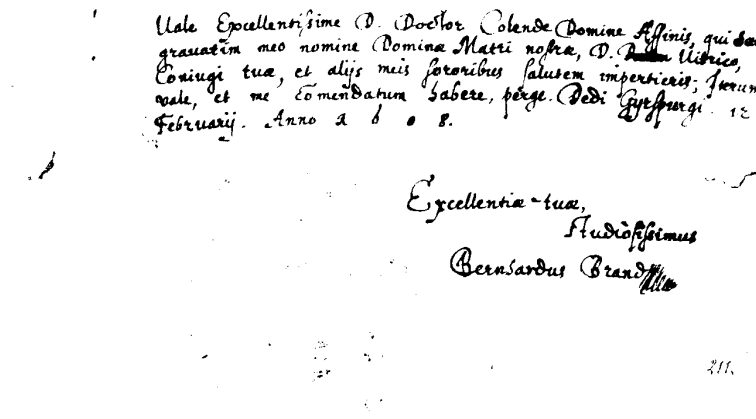
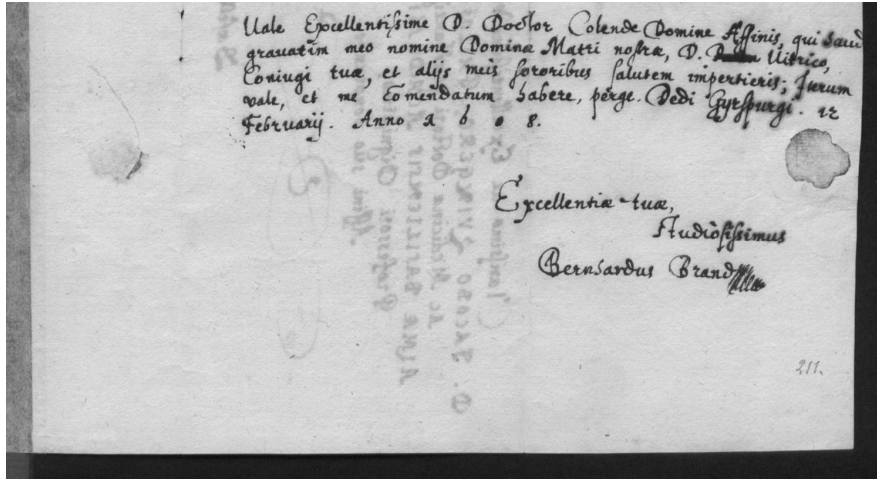


Рис. 8. Приклад бінарizaції зображення документа з сайту DIBCO

Результат оброблення можна зіставити з результатами оброблення такого ж зображення на сайті DIBCO [10].

Відмітимо, що правильно підібрана тестова вибірка надає можливості роботи з достатньо широким класом документів без зміни меж діапазонів відношення сигнал/шум. Наприклад, для обох зображень (рис. 7, 8) використано їхнє однакове налаштування.

Висновки

Запропоновано спосіб бінарizaції зображень документів, який ґрунтується на використанні вейвлет-аналізу. Також запропоновано спосіб додаткового оброблення з використанням класифікації за відношенням сигнал/шум. На цій основі

розроблено відповідне ПЗ. Результати тестування з використанням ПЗ показали, що запропонований спосіб забезпечує стабільні результати бінаризації для доволі широкого за якісними ознаками класу зображень документів, включаючи зображення зі слабким контрастом і суттєвими ознаками згасання.

Практична значущість роботи в тому, що розроблене ПЗ входить до складу програмного пакету з підготовки зображень документів до мікрофільмування з використанням комп'ютерних технологій.

1. Японцы помогут оцифровать манускрипты Ватикана. URL: https://www.bbc.com/russian/society/2014/03/140321_vatican_library_digital
2. Путинцев Д.Н., Кац В.А., Арлазаров Н.В. Локализация открытой воды и разводий на изображениях поверхности Арктики. *Международный журнал прикладных и фундаментальных исследований*. 2017. № 1. С. 341–343.
3. Артюхова О.А., Самородов А.В. Разработка алгоритма автоматической сегментации флуоресцентно-микроскопических изображений препаратов клеточных культур для задач микробиологии. *Наука и образование*. 2013. № 6. С. 275–288.
4. Горитов А.Н. Предварительная обработка изображений в системах технического зрения. *Доклады ТУСУР*. 2018. Т. 21, № 4–1. С. 53–58.
5. Шарий Т.В., Лялин Р.О., Гукай А.Е., Котенко В.Н. Распознавание текста в зашумленных изображениях сканированных документов. *Вестник Донецкого национального университета*. Сер. А: Естественные науки. 2016. № 2. С. 91–98.
6. Wojciech Bieniecki. Multi-pass approach to adaptive thresholding based image segmentation // Wojciech Bieniecki, Szymon Grabowski. Proc. of the 8th International IEEE Conference CADSM 2005. Lviv-Polyana, 2005. P. 418–423.
7. Федоров А. Бинаризация черно-белых изображений: состояние и перспективы развития URL: <http://it-claim.ru/Library/Books/ITS/wwwbook/ist4b/its4/fyodorov.htm>
8. Вдовин В.А., Муравьев А.В., Певзнер А.А. Метод адаптивной бинаризации растрового изображения. *Ярославский педагогический вестник*. 2012. Том III (Естественные науки). № 4. С. 65–69.
9. Ioannis Pratikakis, Konstantinos Zagoris, Panagiotis Kaddas, Basilis Gatos. ICFHR 2018 Competition on Handwritten Document Image Binarization (H-DIBCO 2018). 16th International Conference on Frontiers in Handwriting Recognition. 2018. P. 489–493.
10. DIBCO 2019. IDCAR 2019 Document Image Binarization Competition. URL: <https://vc.ee.duth.gr/dibco2019>
11. Астафьева Н.М. Вейвлет-анализ: основы теории и примеры применения. *УФН*. 1996. Т. 166. № 11. С. 1145.
12. Дьяконов В. МАТЛАВ. Обработка сигналов и изображений. Спец. справ. Санкт-Петербург: Питер, 2002. 397 с.
13. Лоза В.Н., Ленков Е.С. Особенности применения пакетных алгоритмов вейвлет-анализа при обработке сигналов. *Системы обробки інформації*. 2016. № 7(144). С. 66–71.
14. Любушин А.А. Программа WaveSpectExp оценки эволюции вейвлетной спектральной экспоненты для группы временных рядов в скользящем временном окне. Описание метода. URL: http://old.ifz.ru/scient_directions/data_monitoring/09/WaveSpectExp.pdf
15. Єгоров П.Н., Петренко В.И., Шумейко Н.А., Яценко С.Я. Применением методов вейвлет-анализа для контрастирования рентгенографических изображений. Матеріали 18 Міжнар. наук.-техн. конф. «Леотест-2012». Львів, 2012. С. 183–188.
16. Добеши И. Десять лекций по вейвлетам. Ижевск, 2001. 464 с.
17. http://wiki.technicalvision.ru/index.php/Корреляционное_сопоставление_изображений
18. Єгоров П.М. Нечітка система визначення якості цифрових зображень документів, які надано на мікрофільмування. *Реєстрація, зберігання і оброб. даних*. 2016. Т. 18. № 4. С. 67–78. <https://doi.org/10.35681/1560-9189.2016.18.4.104133>

Надійшла до редакції 04.08.2020