

Стиск зображень на основі виділення і кодування об'єктів з різною візуальною якістю

В.Г. Іванов, М.Г. Любарський, Ю.В. Ломоносов

Національна юридична академія України імені Ярослава Мудрого, Харків
nuau@bestnet.kharkov.ua

Анотація

У даній роботі приводяться практичні результати комп'ютерного моделювання по стиску деяких класів зображень. Зображення, перед кодуванням, попередньо розбивається на дві складові – об'єкт і тло. Поділ зображення здійснюється в площині вейвлет-коефіцієнтів. Кодування кожної складової зображення виробляється окремо з різним ступенем точності на основі класичної схеми JPEG-формату.

1. Вступ

В останні роки з'явилися роботи зі стиску зображень, у яких знайшли відображення й одержали подальший розвиток методи стиску другого покоління, що використовують ефективну модель "контур + текстура", а також методи сегментації зображень [1, 2, 3].

Результати цих досліджень, з використанням вейвлет-перетворень і автоматичної класифікації дозволили створити на практиці найсучасніший на сьогоднішній день графічний формат DjVu-компресії, основна ідея якого полягає у відділенні тексту (переднього плану) від тла (заднього плану) на зображенні й у використанні для кожного з них найбільш придатного методу кодування [4].

Однак застосований у форматі метод мультимасштабної кластеризації не завжди правильно розділяє зображення на об'єкт і тло, а при кодуванні напівтонових реалістичних зображень, наприклад, портрет він узагалі перестає працювати і зводиться до традиційного алгоритму вейвлет-копресії JPEG 2000. Тому дуже актуальною є задача розробки і дослідження нових схем скорочення статистичної, структурної і психофізичної

надмірності зображень на основі об'єднання і комбінування незалежних методів стиску даних і методів розпізнавання образів, що дозволить розглядати ці методи з єдиних позицій і вирішити задачі гранично стиснутого опису повідомлень.

2. Виділення об'єктів на площині зображень

Ідея кодування зображення з різним ступенем якості реалізована в графічному форматі DjVu [4], призначеному для ефективного збереження змішаних зображень, що вміщують у собі текст і ілюстрації. Зображення розділяється на передній та задній плани і інформація про які утримується у бітовій площині. Поділ зображення на передній план (у нашому випадку – аналог об'єкта) і задній план (тло) засновано на перебуванні границь різких перепадів яскравості між областями з рівномірним розподілом цього параметру. Кодування переднього і заднього плану зображення виробляється на основі вейвлет-перетворення даних, отриманих з різним дозволом. При цьому дозвіл заднього плану в кілька разів менше ніж дозвіл переднього плану. Таким чином, оброблюване зображення у форматі DjVu кодується з різним ступенем якості, що в порівнянні з форматом JPEG-2000 (JP2), де також використовується вейвлет-перетворення, але для всієї площини зображення, має визначену перевагу в стиску на деяких класах зображень.

Обмеженням приведеного методу виділення переднього плану є те, що значимий об'єкт повинен мати рівномірний розподіл яскравості усередині себе, бути замкнутим за формою і мати чіткі межі

перепаду яскравості з заднім планом зображення. У випадку, коли найбільш інформативна область зображення не має замкнутої форми з постійним значенням яскравості, цей алгоритм не робить поділу зображення і кодує його з рівним ступенем якості.

Головним питанням даної роботи є реалізація алгоритму виділення найбільш інформативних областей зображення, таких що не мають рівномірного розподілу яскравості і замкнутих меж. У якості такого зображення, з бібліотеки стандартних зображень <http://www.icsl.ucla.edu>, був обраний файл zelda.bmp, що представлений на рис. 1. Параметри вихідного зображення: розмір - 256x256, глибина кольору в градації сірого - 8 біт. Вейвлет коефіцієнти вихідного зображення після одного кроку перетворення представлені на рис. 2.

На практиці реалізація вейвлет-перетворень зводиться до застосування біортогональних вейвлет-базисів, у яких формула розкладання (1) і формула синтезу (2) є основними [5].

Якщо $\{a_i\}$ – коефіцієнти розкладання залишкового члена, то $\{b_i\}$ і $\{c_i\}$ – коефіцієнти відповідно нового залишкового члена

$$b_i = \sum_s h_s a_{2i+s} \quad \text{и.} \quad c_i = \sum_s g_s a_{2i+s} \quad (1)$$

Аналогічно, якщо відомі коефіцієнти $\{b_i\}$ і $\{c_i\}$, то коефіцієнти $\{a_i\}$ можна відновити по формулі

$$a_i = \sum_s h_s b_{i-s} + \sum_s g_s c_{i-s} \quad (2)$$

Для розкладання і відновлення сигналу використовуються дві різні пари квадратурних дзеркальних фільтрів: пара $\{\tilde{h}, \tilde{g}\}$ при розкладанні і пара $\{h, g\}$ при відновленні, де $\{g_i\}$ послідовність, сполучена до послідовності $\{h_i\}$, так, що $g_i = (-1)^i h_{1-i}$.

Основне рівняння, у цьому випадку має вигляд

$$\overline{h(\omega)} \tilde{h}(\omega) + \overline{h(\omega + \pi)} \tilde{h}(\omega + \pi) = 2$$

Для цього рівняння існує сімейство рішень:

$$\tilde{h}_{n,m}(\omega) = \sqrt{2} \left(\frac{1 + e^{i\omega}}{2} \right)^n ;$$

$$h_{n,m}(\omega) = \sqrt{2} \left(\frac{1 + e^{i\omega}}{2} \right)^2 P_m \left(\sin^2 \frac{\omega}{2} \right) e^{-im\omega} \quad (3)$$

де $P_m(x) = \sum_{s=0}^{m-1} C_{m-1+s}^s x^s$ – багаточлен Дебеші.

Для вейвлет перетворення вихідного зображення у виразі (3) використовувалася пара квадратурних дзеркальних фільтрів h і \tilde{h} з параметрами $n=1$ і $m=2$, що представлені на рис. 3.



Рис.1. Вихідне зображення.

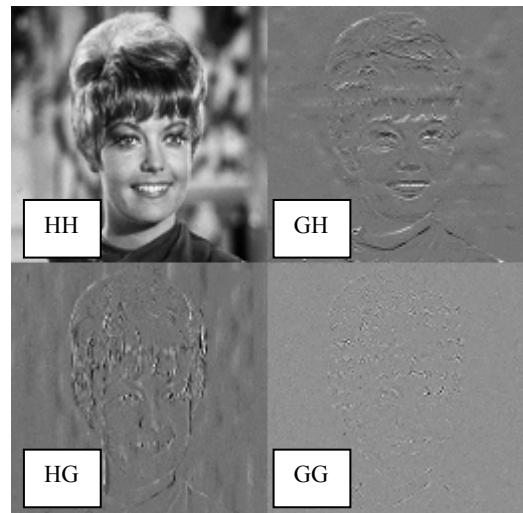


Рис. 2. Вейвлет коефіцієнти після одного кроку перетворення.

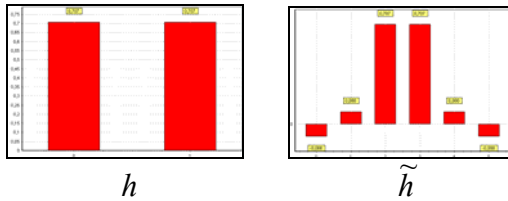


Рис. 3. Пари квадратурних дзеркальних фільтрів.

При побудові поділяючої бітової площини, використовувалися вейвлет коефіцієнти HG і GH областей, отримані згідно виразу (1). Після усередненого підсумування заданих вейвлет коефіцієнтів $(HG+GH)/2$ одержимо зображення на рис.4.

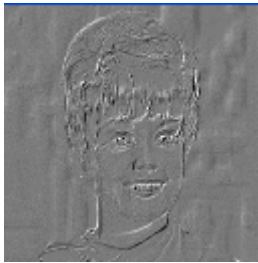


Рис. 4. Зображення, отримане в результаті усередненого підсумування вейвлет коефіцієнтів.

Це зображення має більш рівномірний розподіл яскравості в порівнянні з оригіналом (рис. 1) чи залишковим членом розкладання НН (рис. 2). Останнє досягається завдяки застосуванню відповідної пари фільтрів при вейвлет перетворенні. Вихідне зображення розкладається на “грубу” (НН) і “тонку”(GG) структуру. Області HG і GH мають більш гладку форму яскравості на всій площині, але при цьому зберігають найбільш істотні її перепади.

У результаті граничної обробки зображення (див. рис. 4), одержимо сукупність крапок, різниця між якими перевищує граничне значення $(|X_i - X_{i+1}| \geq KX_{cp})$ (див. рис. 5а). Граничне значення прив'язувалося до середнього значення яскравості X_{cp} і обиралося таким чином, щоб виключити наявність випадкових крапок, тобто тих крапок, що не мали у своєму найближчому оточенні собі подібних. На рис. 5б, представлені

граничні крапки, що розділяють область об'єкта і тла. Після з'єднання граничних крапок між собою одержимо бітову площину що розділяє об'єкт і тло (див. рис. 5в). На рис. 5г представлена бітова площина поділу зображення після сегментної розбивки на фрагменти 4x4.

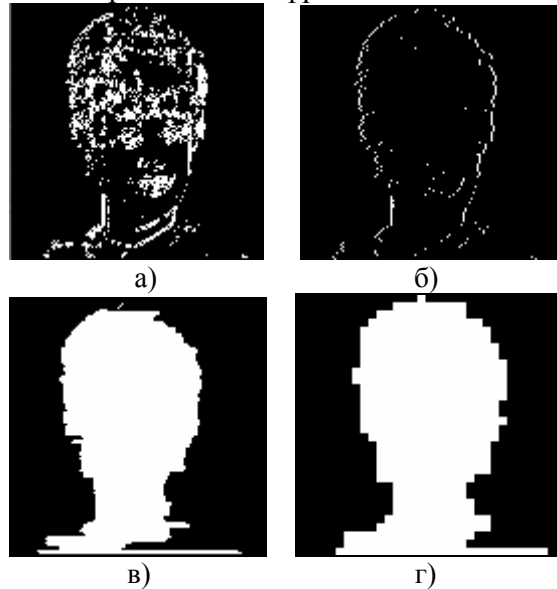


Рис. 5. а) Сукупність крапок, які перевищили різницевий поріг; б) граничні крапки поділу об'єкта і тла; в) бітова площина поділу; г) бітова площина після сегментації.

3. JPEG-компресія зображень з різним ступенем точності.

У запропонованому методі по поділу зображення, об'єктом будемо називати область, що містить крапки, які перевищили поріг, а тлом – область у якій немає різких перепадів яскравості.

Вихідне зображення (рис. 1) так само необхідно розбити на сегменти розміром 8x8, що цілком узгоджується з розміром сегментації бітової площини (рис. 5, г), і здійснити кодування об'єкта і тла зображення у класичному стандарті JPEG-технологій.

Різний ступінь точності кодування об'єкта і тла визначається кількістю косінусних коефіцієнтів, що використовуються в робочих матрицях об'єкта і тла. Для тла виберемо мінімальне значення -1 , а для об'єкта максимальне -64

коефіцієнта. На рис. 6 приведена залежність загального коефіцієнта стиску всього зображення ($K_{ст}$) від середньоквадратичної помилки (СКП) на площині об'єкта.

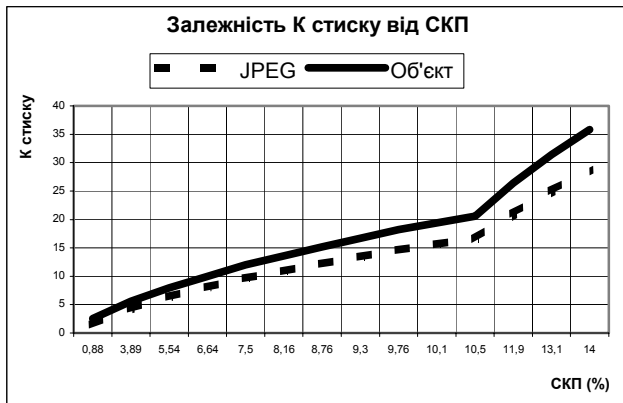


Рис. 6. Залежність загального коефіцієнта стиску зображення ($K_{ст}$) від середньоквадратичної помилки (СКП) на площині об'єкта



а) $K_{стиску} = 26,4$ СКП=12%.



б) $K_{стиску} = 21$, СКП=12%

Рис. 7. Зображення після кодуювання.

4. Висновки

Запропонований метод поділу зображення на об'єкт і тло дозволяє кодувати їх роздільно. Наприклад, з різним ступенем якості чи навіть різними методами. Це дає можливість закодувати найбільш інформативні області зображення з більш високою якістю, чим всі інші його частини. Завдяки цьому, зберігши високу якість інформативних областей (об'єктів), можна підвищити загальний для всього зображення коефіцієнт стиску. Так для зображення zelda.bmp коефіцієнт стиску в JPEG-форматі збільшується з 21 до 26,4, рази завдяки більш сильній компресії тла при однаковій помилці відновлення (рис. 7).

5. Література

1. Кунт М., Икономопулос А., Кошер М. Методы кодирования изображений второго поколения // ТИИЭР. – 1985. – Т. 73, № 4. – С. 59-86.
2. Gupta Maya R., Stroilov A. Segmenting for wavelet compression // DCC. – 2005. (<http://www.computer.org/>).
3. Иванов В.Г. Кодирование изображений на основе автоматической классификации и позиционирования фрагментов // Матеріали 12-ї Міжнародної конференції з автоматичного управління (Автоматика – 2005). Харків: НТУ “ХПІ”. – 2005. – Т. 3. – С. 80-81.
4. <http://www.dejavu.research.att.com/djvu/sci/djvuspec>.
5. Иванов В.Г., Любарский М.Г., Ломоносов Ю.В. Фурье и вейвлет-анализ изображений в плоскости JPEG технологий // Проблемы управления и информатики. – 2004. - № 5. – С. 111-124.