УДК 519.677
С.А. ТАЯНОВ

Національний університет “Львівська політехніка”,
кафедра автоматизації та комплексної механізації машинобудівної промисловості

ДОСЛІДЖЕННЯ КОМПРЕСІЇ НА ОСНОВІ КОМПОНЕНТНОГО АНАЛІЗУ КЛАСТЕРІВ ЗОБРАЖЕНЬ ШТУЧНОГО ЗОРУ РОБОТА

Досліджено ефективність стискання зображень штучного зору робота на основі розбиття зображення на окремі кластери та подальшої компресії за допомогою компонентного аналізу (перетворення Корунена-Лоєва) .

Efficiency of image compression of robot artificial vision on the base of cutting image on separate clusters and further compression with help of component analysis (Hetteling or Korhunen-Loeve transformation) is investigated

Вступ. Збереження графічної інформації в пам’яті робота вимагає великого обсягу пам’яті, Тому актуальною є компресія отриманої графічної інформації з давачів штучного зору і збереженні її в такому вигляді на носіях. Крім того, відновлення стиснутого зображення має відбуватися в режимі реального часу, тому кількість математичних операцій має бути зведена до мінімуму.

Аналіз останніх досліджень. Відомі ефективні алгоритми для компресії зображень на основі двовимірного дискретного косинусного перетворення (2DCT) [1–3], які дозволяють отримувати високий ступінь стискання зображень за рахунок того, що в матриці частотних коефіцієнтів, яка втримується з вихідної матриці зображення після дискретного косинусного перетворення, низькочастотні коефіцієнти розташовані ближче до лівого верхнього краю, а високочастотні внизу праворуч. Це є важливо, оскільки більшість зображень складається з низькочастотної інформації, а високочастотні компоненти матриці можна відкинути. Цей принцип реалізований в графічному форматі JPEG [4]. Також відома серія алгоритмів на основі хвильового перетворення (Wavelet transformation), які знайшли використання в стандарті графічного формату JPEG2000 [5]. Вищенаведені алгоритми є універсальними і можуть бути застосовані для будь-якого типу зображень. Перетворення за цими алгоритмами виконується за допомогою певних матриць перетворення, які є незмінними. Це означає, що для конкретного зображення матриця перетворення не завжди буде оптимальною, а отже ми не досягнемо максимальної компресії зображення. Тому актуальним є розробка адаптивних алгоритмів для конкретного зображення або серії зображень, що дозволить істотно збільшити коефіцієнт компресії.

Постановка задачі. У статті пропонується адаптивний алгоритм компресії монохромних зображень на основі розбиття зображення на кластери та подальшої компресії за допомогою перетворення Корунена-Лоєва. Наводяться результати дослідження запропонованого алгоритму, а також пропонуються подальші алгоритми для покращання компресії монохромних зображень.

Виклад основного матеріалу. Застосування перетворення Корунена-Лоєва для компресії зображень є звичайно неефективним на практиці внаслідок великої кількості обчислень. Крім того, велика матриця перетворень зображення (матриця власних векторів) має бути збережена разом зі стиснутими даними, що унеможливлює компресію. Але перетворення Корунена-Лоєва має властивість мінімізації середньоквадратичної похибки при використанні лише кінцевої кількості базисних функцій у розкладі. Ця властивість полягає у тому, що вона гарантує неможливість отримання меншої у середньоквадратичному змісті помилки апроксимації за допомогою іншого розкладу. Крім того, використання цього перетворення добре показало себе для компресії аудіоінформації [6].

Дискретне перетворення Корунена-Лоєва. Перетворення Корунена-Лоєва основане на матриці перетворення, яка (на відміну від косинусного перетворення, в якому вона наперед визначена) визначається на основі цих вибірок. Ця матриця має властивість, що вона декорелює коваріаційну матрицю, це здебільшого зумовлює сильну компресію.

Розглянемо дискретне перетворення Корунена-Лоєва так як неперервне перетворення не має практичного застосування при роботі з зображеннями.

Якщо x_i – i-та вибірка значень, то перетворення Корунена-Лоєва можна записати як

$$x_i = \sum_{j=1}^{n} c_{ij}\phi_j \ , \tag{1}$$

де c_i – вектор коефіцієнтів перетворення Корунена-Лоєва для i-ї вибірки:

$$c_i = \begin{bmatrix} c_{i1} \\ c_{i2} \\ \dots \\ c_{in} \end{bmatrix}, \tag{2}$$

ϕ – матриця власних векторів $\phi = (\phi_1, \phi_2, \phi_3, ..\phi_n)$; n – розмір вибірки.

Вираз [1] – можна записати в матричній формі:

$$x_i = c_i \cdot \phi \tag{3}$$

Матриця власних векторів визначається з коваріаційної матриці C, яку одержують так:

$$Cov = \frac{1}{M}\sum_{i=1}^{M}(x_i - \overline{x})(x_i - \overline{x})^T \ , \tag{4}$$

де M – кількість вибірок; $\overline{x}$ – середнє значення для всіх вибірок, тобто $\sum_{i=1}^{M}(x_i - \overline{x}) = 0$.

З виразу (3) можна знайти коефіцієнти перетворення

$$c_i = x_i \cdot \phi^{-1}. \tag{5}$$

Методика компресії зображень за допомогою ДПКЛ. Для зменшення кількості обчислень та розміру матриці власних векторів розбиваємо матрицю зображення на кластери.

Нехай маємо зображення розміром $m \times n$. Розбиваємо цю матрицю на прямокутні кластери розміром 8×8, кластери нумеруємо зліва направо та згори донизу. Розмір кластера 8×8 використовується в алгоритмі JPEG [4] і добре себе зарекомендував. Нумеруємо елемент кожного кластера зліва-направо та згори–донизу рис. 1. Тобто лівий верхній елемент i-го кластера буде позначатися як S_{i1}, а правий нижній елемент позначатиметься як S_{i64}.

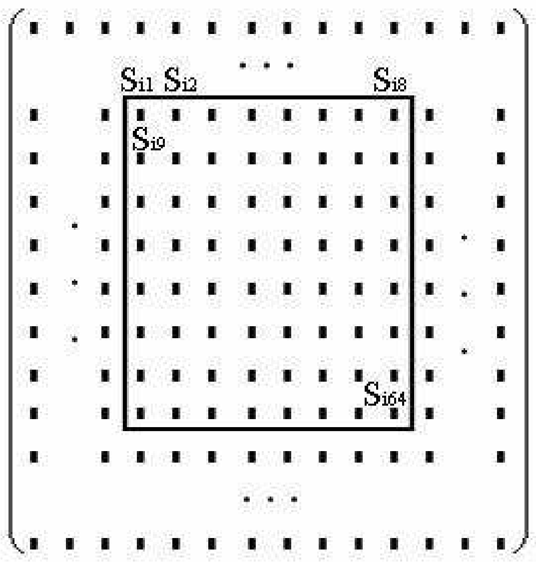

Рис. 1. Розбиття матриці зображення
на прямокутні кластери

В результаті отримуємо $\frac{m \times n}{8 \times 8}$ векторів кластерів з 64 елементів кожний. Якщо S_i – i-й кластер, то S_{ij} – j-й елемент i-го кластера (j=1,2,..64); A – матриця зображення; A_{ij} – елемент матриці A, тоді коваріаційна матриця для всіх кластерів S виглядатиме:

$$Cov = \frac{8 \times 8}{m \times n} \sum_{i=1}^{\frac{m \times n}{8 \times 8}} (S_i - \bar{x})(S_i - \bar{x})^T \tag{6}$$

де $\bar{x} = \frac{1}{m \cdot n} \sum_{i=1}^{m} \sum_{i=1}^{n} A_{ij}$.

З коваріаційної матриці отримуємо матрицю власних векторів

$$\varphi = eigenvector(Cov) \quad (7),$$

яка складається з 64 власних векторів, кожний з яких складається з 64 елементів. Варто зауважити, що при збереженні матриці на кожний елемент виділяється лише 1 байт. Тому перед збереженням матрицю необхідно нормалізувати та квантувати. Тобто кожний елемент матриці власних векторів має бути цілим числом в діапазоні від -127 до +127. Це можна записати як:

$$\varphi^{norm} = ENT\left[\frac{127}{|\varphi_{\max}|} \cdot \varphi\right], \tag{8}$$

де $\varphi_{\max}$ – максимальний елемент матриці φ, ENT – найближче ціле число.

Знаходимо вектори коефіцієнтів перетворення Корунена-Лоєва для кожного кластера з виразу (5). Якщо всі кластери S_i подати у вигляді матриці $S = (S_1, S_2, S_3, .., S_{\frac{m \times n}{8 \times 8}})$, то можна записати

$$c = S \cdot (\varphi^{norm})^{-1} . \tag{9}$$

У результаті отримаємо матрицю коефіцієнтів перетворення Корунена-Лоєва розміром $(\frac{m\times n}{8\times 8})\times 64$. Приблизний вигляд матриці коефіцієнтів для картинки 256 х 256 зображений на рис. 2. На рис. 2 c_{ij} – j-й коефіцієнт для i-го кластера.

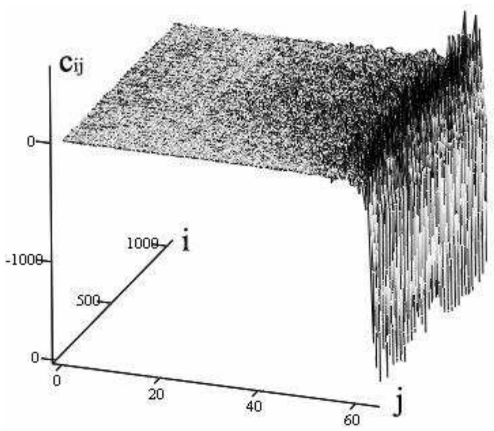

Рис. 2. Матриця коефіцієнтів перетворення

Матрицю коефіцієнтів перетворення так само необхідно нормалізувати, тобто

$$c^{norm}=\frac{127}{|c_{\max}|}\cdot c\,. \qquad (10)$$

Якщо в матриці коефіцієнтів перетворення c виділити тільки найістотніші значення за амплітудою для кожного вектора власних значень c_i (наприклад ті в яких j>40–45 (рис. 2) і тільки їх зберігати і враховувати, то можна зменшити кількість інформації, що зберігається. Крім того, необхідно зберігати тільки значення в матриці власних векторів, які відповідають збереженим коефіцієнтам перетворення. Для цього випадку коефіцієнт компресії можна порахувати за формулою

$$K=\frac{m\cdot n}{N(\frac{m\cdot n}{64}+64)}, \qquad (11)$$

де N – кількість найістотніших значень у кожного вектора коефіцієнтів перетворення, які враховуються. Залежність співвідношення сигнал-шум SNR від N зображено на рис 3:

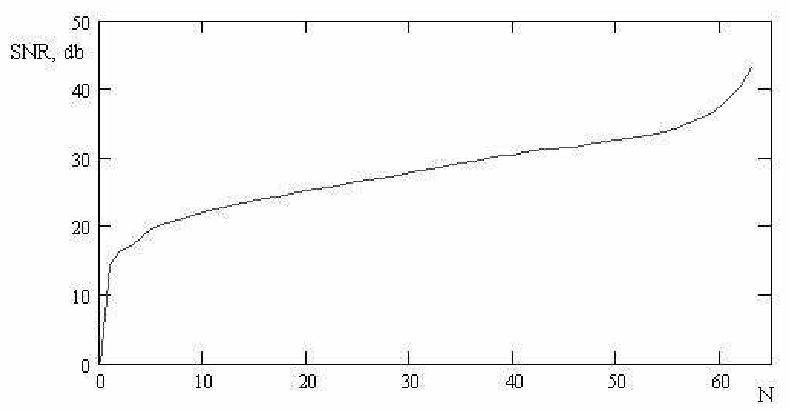

Рис. 3. Залежність SNR від кількості відліків у кожному векторі коефіцієнтів перетворення N

З рис. 3 очевидно, що SNR зростає стрімко при кількості врахованих значень N від 0 до 8 ш при N=8 SNR $\approx$ 20db , у разі подальшого зростання N SNR зростає повільніше. Тому для досягнення найефективніших параметрів компресії необхідно вибирати N від 8..10 до 40, де SNR $\approx$ 30db. Більше ніж 40 N брати не доцільно, оскільки отримуємо за формулою (11) коефіцієнт компресії, трохи більший ніж 1.

Залежність SNR від коефіцієнта компресії зображена на рис. 4.

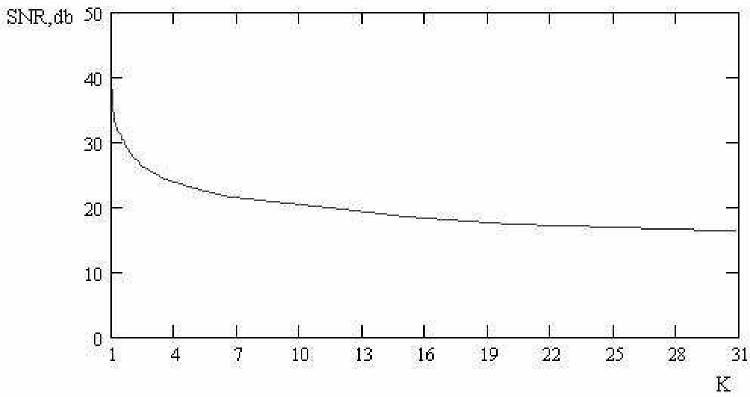

Рис. 4. Залежність SNR від коефіцієнту компресії K

З рис. 4 очевидно, що SNR швидко спадає зі зростанням коефіцієнта компресії K від 0 до 4. Тому для зображень, для яких вимагається висока якість відновлення SNR>25 можливо досягти коефіцієнт компресії порядку K=1.5..4. Для зображень, для яких SNR > 20 (наприклад зображення великих об'єктів в пам'яті робота), можна досягнути коефіцієнта компресії K=10.

Для відновлення зображення необхідно відновити кластери цього зображення. Для отримання кожного кластера необхідно матрицю власних векторів перемножити з вектором коефіцієнтів перетворення Корунена-Лоєва для цього кластера (3). Після того скласти зображення з окремих кластерів.

Проводилась компресія матриці власних векторів за допомогою двовимірного косинусного перетворення, яка показала, що можна додатково досягнути збільшення коефіцієнта компресії на 10–15 % . Але це доцільно проводити тільки за малих коефіцієнтів компресії.

Висновки. Отже, виконані дослідження ефективності компресії зображень штучного зору робота за допомогою розбиття зображення на кластери та подальшого компонентного аналізу. Дослідження показали, що для зображень великих об'єктів можна досягти коефіцієнта компресії K=10.

1. Said A. and Pearhman W.A. A new fast and efficient image codec based on set portioning in hierachical trees // IEEE Trans. On Circuts and Systems for Video technology. – June 1996. – Vol. 6, no. 3. – P. 243–250. 2. Xiong Z., Guleryuz O. and Orchard M.T. A DCT-based Embedded Image Coder // IEEE Signal Processing Letters. – October, 1996. 3. Таянов С.А. Підвищення коефіцієнта стиску монохромних зображень на основі косинусного перетворення // Вісн. Нац. ун-ту "Львівська політехніка". – 2004. – № 515: Оптимізація виробничих процесів і технічний контроль у машинобудуванні та приладобудуванні. – С. 135–137. 4. JPEG: ITU-T Rec. T.81-ISO/IEC. No. 10918-1, "Information Technology – Digital compression and Coding og Continuous-Tone Still Images", 1993. 5. Taubman D.S. and Marcellin M.W. JPEG2000: Fundamentals, Standards and Practice. Kluwer Academic Publishers, Boston, 2002. 6. Sergey Tayanov, Vitalij Tayanov. The error analysis of speech signal computer compression by Karhunen-Loeve transform // Proceedings of the XI Polish-Ukrainian Conference on "CAD in Machinery Design. Implementation and Educational Problems". – 2003. – С. 113–119.