

## УЩІЛЬНЕННЯ ЗОБРАЖЕНЬ З РЕГУЛЮВАННЯМ ВТРАТ НА ОСНОВІ НЕЙРОПОДІБНИХ СТРУКТУР ГЕОМЕТРИЧНИХ ПЕРЕТВОРЕНЬ

© Поліщук У., 2010

Подано основи розробленого методу ущільнення півтонових та кольорових зображень на основі нейроподібних структур моделі геометричних перетворень, що забезпечує представлення скомпресованих даних у форматі чисел з фіксованою комою. Вперше для нейромережних методів ущільнення зображень реалізовано принципову можливість компресії без втрат або з регульованими втратами.

**Ключові слова:** автоасоціативна нейронна мережа, модель геометричних перетворень, компресія зображень, обчислення у форматі з фіксованою комою.

**This paper describes a basis of the developed color and gray-scale images compression method using neurolike Geometrical Transformation Machine structure, which provides representation of the compressed data in a fixed-point format. Principal possibility to make compression using auto-associated neural networks without or with management information losses has been realized for the first time.**

**Keywords:** auto-associated neural networks, geometrical transformation machine structure, image compression, fixed point computation.

### Вступ

Виникнення і розвиток технологій цифрових зображень стимулювали одночасну появу ефективних методів їх ущільнення. Окрім існуючих методів компресії зображень, представлених серіями стандартів JPEG, MPEG та H.26x, нові технології, побудовані на нейронних мережах та генетичних алгоритмах мають відповідні шанси у майбутньому знайти свою нішу у сфері обробки зображень.

Очевидно, що цифрові зображення належать до інформаційно насичених типів даних, отже, потребують багато місця в пам'яті для зберігання та затрат часу під час передавання каналами зв'язку. Використовуючи різноманітні техніки ущільнення зображень, вдається зменшити надлишкову інформацію, яка зберігається у зображеннях. У результаті зменшуються розміри файлів зображення, час передавання [5] та місце в пам'яті комп'ютера.

Штучні нейронні мережі (ШНМ), які забезпечують компресію за рахунок властивостей узагальнення (генералізації), надають можливість отримувати значні коефіцієнти компресії, демонструючи переваги над існуючими традиційними методами ущільнення [1].

Найпопулярнішими нейропарадигмами стосовно до задач ущільнення даних вважають багатопарові перцептрони в автоасоціативному режимі, самоорганізаційні карти Когонена, мережі ART та мережі радіальних базових функцій (RBF) [2, 3, 4].

Метою цієї роботи є подальший розвиток і вдосконалення нейромережних методів ущільнення зображень на основі застосування високоефективних нейроподібних структур геометричних перетворень (ГП).

Проаналізуємо особливості відомого нейромережного методу ущільнення, що реалізується на основі багатопарових перцептронів в автоасоціативному режимі з використанням так званого принципу «звуженого горла».

Автоасоціативні нейронні мережі такого типу (рис.1) забезпечують компресію зображень шляхом трансляції векторів даних великої розмірності із входу на вихід мережі через порівняно «вузький» канал, оскільки кількість нейронів у прихованому шарі менша ніж кількість нейронів на вході та виході. Тобто в таких структурах відбуваються послідовні перетворення: з простору

великої розмірності → в простір зменшеної розмірності → в простір початкової розмірності. Оскільки частина координат проміжного простору перетворюється відкидається, ущільнення даних відбувається принципово з втратами. В процесі роботи алгоритм зворотного поширення похибки мінімізує похибку. Це означає, що ваги зв'язків вхідного та прихованого шару працюють на власне компресію даних, тоді як прихований та вихідний шари здійснюють декомпресію. Сигнали на виходах нейронних елементів прихованого шару відповідають головним компонентам перетворення (Principal Components Analysis – PCA), тобто саме вони формують інформаційне ядро. Чим менша кількість нейронів у прихованому шарі тим вищий коефіцієнт ущільнення, однак при цьому знижується якість відтвореного зображення [9].

До недоліків цього методу належить згадана вище можливість здійснювати ущільнення лише з втратами інформації та значні часові затримки на виконання процедури навчання мережі.

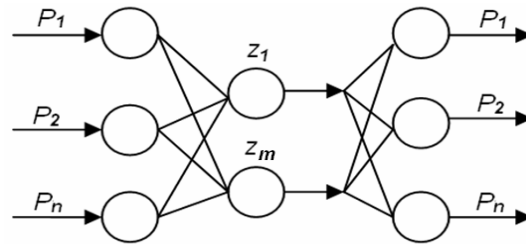


Рис. 1. Автоасоціативна нейронна мережа з «вузьким горлом»

Відомі також застосування для мети ущільнення даних нейроподібних структур геометричних перетворень, які, однак, не передбачають компресії зображень без втрат інформації або з регулюванням рівня таких втрат [8]. Тобто ефекту ущільнення в них досягають також шляхом відкидання певного числа НЕ прихованого шару мережі.

Розглянемо запропоновані у роботі нові метод і алгоритм ущільнення даних, що ґрунтуються на використанні особливостей кодування інформації про зображення.

### Процедура ущільнення зображень

Розглянемо варіант розбиття зображення на блоки розміром  $N$  пікселів, (наприклад,  $N=64$ ,  $8 \times 8$  пікселів). Беручи до уваги відповідність вектора вхід-вихід мережі блокових зображень, використовуємо 3-шарову нейромережу на 64 входи та 64 виходи, блоки зображення на входи при навчанні і застосуванні подаються послідовно. Якщо зображення напівтонове, пікселі набувають одного із значень інтенсивності будь-якого одного кольору, діапазон яскравості якого задається від 0–255 у 8-бітній розрядній сітці із фіксованою комою. Це один із найпоширеніших типів представлення зображень. У випадку кольорового зображення кожен піксель кодується 24-бітовим числом, де компоненти червоного, зеленого та голубого кольорів, згідно з моделлю RGB, представлені у 8-бітній розрядній сітці. Тобто для кольорового зображення використовується 192 ( $64 \times 3$ ) відповідно входів та виходів мережі, де кожен піксель зображення представлений трьома компонентами вектора.

Подамо матрицю реалізацій для нейромережі у вигляді (1). Кожен рядок матриці відповідає вектору-реалізацій, тобто відповідному блоку зображення, число яких дорівнює  $M$ , а число стовпців дорівнює кількості пікселів зображення у блоці  $N$  (для кольорового зображення число стовпців матриці збільшене втричі).

$$x_{i,j} = \begin{Bmatrix} x_1 & x_2 & \dots & x_N \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \end{Bmatrix}, \text{ де } i = \overline{1, M}, j = \overline{1, N}; \quad (1)$$

Застосування нейронної мережі для будь-яких завдань у загальному випадку передбачає два послідовні етапи навчання і, відповідно, використання навченої мережі. Однак, для задач ущільнення зображень можливими є два варіанти: навчання і застосування мережі на одному і тому самому зображенні; навчання на певному еталонному зображенні і застосування навченої раніше нейромережі в режимі використання для ущільнення і відновлення інших зображень.

### Етап навчання нейронної мережі

Навчання нейронної мережі здійснюється шляхом покрокових геометричних перетворень – трансформацій вхідної матриці, де спосіб розкладу ґрунтується на геометричній моделі і представленні тіла інформаційного об'єкта [6,7]. Кожен вхідний вектор розглядається як точка у  $n$ -вимірному просторі, де  $n$  – кількість компонент вектора. Множина точок-реалізацій формує тіло інформаційного об'єкта (у даному випадку зображення), що моделюється за допомогою нейромережі.

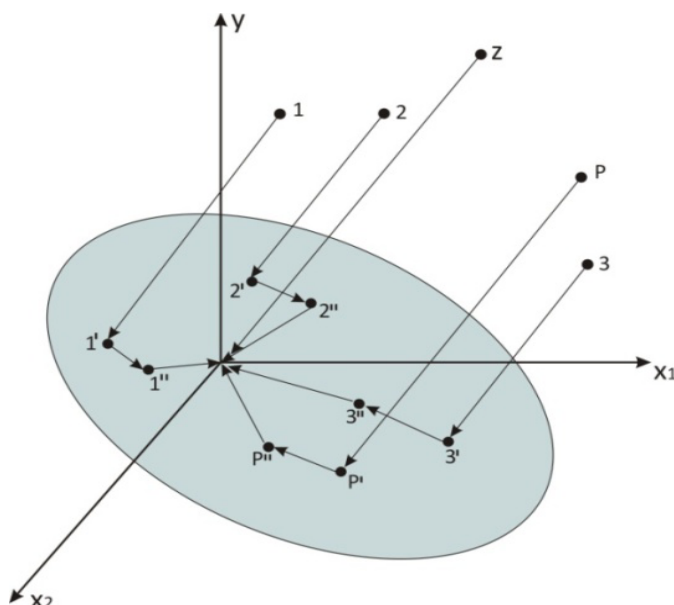


Рис. 2. Проекція точок-реалізацій на нормальну площину

Процедура навчання передбачає такі кроки:

1. Обираємо базовий рядок-вектор, що є найбільш віддаленим від початку координат:

$$xB_j^{(k)} \approx \max_i \left\{ \sum_{j=1}^n (x1_{i,j}^{(k)})^2 \right\}, \quad (2)$$

де  $x_{i,j}^{(k)}$  – вхідна матриця реалізацій;  $k = 0, 1, 2, \dots, k_{\max}$  – номер кроку перетворень.

2. Знаходимо віддаль від кожної  $i$ -ї точки-реалізації до нормальної гіперплощини:

$$K_i^{(k)} = \frac{\sum_{j=1}^n (x_{i,j}^{(k)} \times xB_j^{(k)})}{\sum_{j=1}^n (xB_j^{(k)})^2}, \quad (3)$$

3. Виконуємо  $k$ -й крок ортогональних перетворень – проектування точок на нормальну гіперплощину вимірності  $N-1$ :

$$x_{i,j}^{(k+1)} = x_{i,j}^{(k)} - XB_j^{(k)} \times K_i^{(k)}, \text{ для } i = \overline{1, M}, j = \overline{1, N}, \quad (4)$$

У [6, 7] пропонується обчислювати віддаль, нормовану до максимального значення (3). Описаний підхід забезпечує реалізацію режимів навчання і подальшого використання навченої нейронної мережі. Обчислене так значення коефіцієнта  $K_i^{(k)}$  знаходиться в діапазоні дробових чисел  $[-1; +1]$ , а здійснення перетворень (3) можливе у форматі з плаваючою комою – діапазон чисел чисел в (4) є надто широким. Аналогічний підхід на основі формату представлення чисел з плаваючою комою застосовується і в автоасоціативних багатосарових перцептронах прямого поширення [9].

Очевидно, що подібні представлення не можна вважати оптимальними як в плані збереження ущільнених даних, так і в плані регулювання точності при виконанні компресії.

У роботі запропоновано модифікацію методів навчання та використання нейромережі, яка передбачає обчислення ненормованих віддалей до нормальної площини на основі співвідношення:

$$z_i^{(k)} = \frac{\sum_{j=1}^n (x1_{i,j}^{(k)} \times xB_j^{(k)})}{\sqrt{\sum_{j=1}^n (xB_j^{(k)})^2}} \quad (5)$$

У цьому випадку максимальне значення  $z_i^{(0)} = 255 \times \sqrt{N}$ , де  $N$  – вимірність вектора-реалізації, а мінімальне значення, відмінне від нуля, дорівнює 1. Для наступних  $k = 1, 2, 3, \dots$  відповідні величини зменшуються, тобто необхідну кількість бітів для представлення  $z_i^{(k)}$  нескладно обчислити, а також здійснити наступну мінімізацію цієї кількості для всієї нейромережі. У зв'язку із запропонованою зміною наступні кроки перетворень набувають вигляду:

1. Початковий масив даних, що представляє зображення

$$x_{i,j} = \begin{Bmatrix} x_1 & x_2 & \dots & x_N \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \end{Bmatrix}, \quad \text{де } i = \overline{1, M}, j = \overline{1, N}; \quad (6)$$

2. Обчислюємо центр ваги

$$\tilde{x}c_j = \frac{\sum_{i=1}^m x_{i,j}}{M}, \quad \text{для } j = \overline{1, N}; \quad (7)$$

3. Знаходимо «реальний» (фактичний) рядок, найближчий згідно Евкліда до центру ваги :

$$xc_j \approx \min_i \left\{ \sum_{j=1}^n (\tilde{x}c_j - x_{i,j})^2 \right\}, \quad \text{для } i = \overline{1, M}, j = \overline{1, N}; \quad (8)$$

4. Виконуємо центрування елементів масиву

$$x1_{i,j}^{(k)} = x_{i,j} - xc_j, \quad \text{для } k = 0. \quad (9)$$

Надалі оперуємо центрованими величинами:  $x1_{i,j}^{(k)}$ , для  $i = \overline{1, M}, j = \overline{1, N}$ .

5. Знаходимо базовий рядок (найвіддаленіша точка від початку координат)

$$xB_j^{(k)} \approx \max_i \left\{ \sum_{j=1}^n (x1_{i,j}^{(k)})^2 \right\}. \quad (10)$$

6. Обчислюємо віддаль від початку координат до базової точки

$$D^{(k)} = \sqrt{\sum_{j=2}^n (xB_j^{(k)})^2}. \quad (11)$$

7. Виконуємо нормування базового рядка до одиничного радіуса та обчислюємо коефіцієнт  $z_i^{(k)}$  – віддаль від  $i$ -ї точки до нормальної гіперплощини на  $k$ -му кроці ортогональних перетворень

$$XB_j^{(k)} = \frac{xB_j^{(k)}}{D^{(k)}}, \quad \text{для } j = \overline{1, N} \quad (12)$$

$$z_i^{(k)} = \frac{\sum_{j=1}^n (x1_{i,j}^{(k)} \times XB_j^{(k)})}{D^{(k)}} \quad (13)$$

8. Виконуємо  $k$ -й крок ортогональних перетворень (проекування точок-реалізацій на нормальну гіперплощину):

$$x_{i,j}^{(k+1)} = x_{i,j}^{(k)} - XB_j^{(k)} \times z_i^{(k)}, \quad \text{для } i = \overline{1, M}, j = \overline{1, N} \quad (14)$$

Після кожного кроку ортогональних перетворень вимірність гіперповерхні зменшується на одиницю, а характеристики самого перетворення (коефіцієнти  $z_i^{(k)}$ ) зберігаються в пам'яті як параметри мережі.

Якщо виконується умова:  $x_{i,j}^{(k+1)} \equiv 0$  – усі точки спроектовані у початок координат, навчання мережі завершено; в іншому випадку здійснюємо наступний крок ортогоналізації:  $k = k + 1$ .

#### Етап використання нейронної мережі, відтворення первинних даних

Первинні дані відтворюються в послідовності, зворотній до алгоритму навчання мережі, починаючи із останнього кроку ортогональних перетворень:  $k = k_{\max}, i = 1, M, j = 1, N$ .

Відновлення центрованого та первинного масиву-кластера:

$$x1_{i,j}^{(k)} = x1_{i,j}^{(k+1)} + XB_j^{(k)} \times z_i^{(k)} \quad (15)$$

$$x_{i,j} = x1_{i,j}^{(0)} + xc_j \quad (16)$$

Отже, в результаті на виходах мережі отримуємо масив чисел  $x_{i,j}^{(k)}$ , які відповідають значенням інтенсивності пікселів відновленого зображення.

#### Результати роботи алгоритму

Ефективність розробленого методу оцінювали, порівнюючи характеристики скомпресованого нами зображення за рахунок відкидання нейронів мережі у прихованому шарі з відповідними даними експериментів, отриманими для багатошарових перцептронів традиційного типу [9].

Для апробації методу використовувались півтонові та кольорові зображення. На рис. 3, 4 зображено результат виконання програми "Bitmap coder" на тестових зображеннях «Mandrill» та «Lena\_gray», що кодувались за допомогою нейронної мережі із 64 входами та виходами.

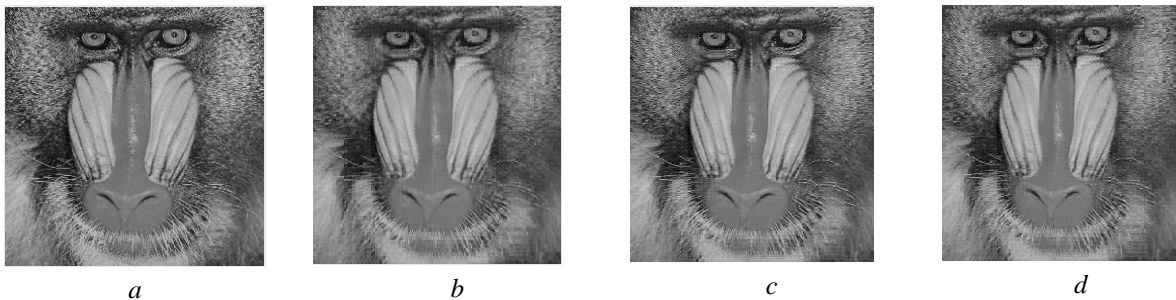


Рис. 3 Результати виконання програми на тестовому зображенні "Mandrill", розмір блоку зображення 8x8: a – початкове зображення; b, c, d – зображення, відновлені після ущільнення із 5, 4 та 3 нейронами відповідно



Рис. 4 Результати виконання програми на тестовому зображенні "Lena\_gray", розмір блоку зображення 16x16: a – початкове зображення; b, c, d – зображення, відновлені після ущільнення із 10, 9 та 4 нейронами відповідно

Розроблений програмний продукт на основі нейроподібних структур ГП в автоасоціативному режимі виконання тестувався на багатьох різноманітних зображеннях. Експерименти проводились із зображеннями розміром 512x512 та розміром блоку 4x4, 8x8 та 16x16 пікселів. Мережа також тестувалась із різною кількістю нейронів у прихованому шарі, в результаті чого ступінь компресії зображень становив від 4:1 до 64:1.

Таблиця 1.

## Результати кодування зображення "Lena\_gray"

Розмір зображення	Розмір блоку зображення	Кількість нейронів прихованого шару	Степінь компресії, CR	Рівень декомпресійних спотворень, PSNR
512 x 512	4 x 4	4	4 : 1	35,9429100026035
512 x 512	8 x 8	5	13 : 1	32,0230442296488
512 x 512	8 x 8	4	16 : 1	30,9674727949574
512 x 512	8 x 8	3	21 : 1	29,3946061270628
512 x 512	16 x 16	10	26 : 1	41,7278964362777
512 x 512	16 x 16	9	28 : 1	40,3227945374787
512 x 512	16 x 16	8	32 : 1	38,9352588043783
512 x 512	16 x 16	4	64 : 1	32,3480579394142
512 x 512	4 x 4	4	4 : 1	35,9429100026035

Таблиця 2.

## Результати кодування зображення "Lena\_color"

Розмір зображення	Розмір блоку зображення	Кількість нейронів прихованого шару	Степінь компресії, CR	Рівень декомпресійних спотворень, PSNR
512 x 512	4 x 4	4	4 : 1	34,6355848854181
512 x 512	8 x 8	5	13 : 1	32,2912822905199
512 x 512	8 x 8	4	16 : 1	31,3306484331729
512 x 512	8 x 8	3	21 : 1	30,0617394832461
512 x 512	16 x 16	20	18 : 1	40,5690528971859
512 x 512	16 x 16	10	26 : 1	35,2594815939445
512 x 512	16 x 16	9	28 : 1	34,6288258727935
512 x 512	16 x 16	8	32 : 1	33,8984868296636
512 x 512	16 x 16	4	64 : 1	29,7584802351575

Таблиця 3.

## Результати кодування зображення "Mandrill\_gray"

Розмір зображення	Розмір блоку зображення	Кількість нейронів прихованого шару	Степінь компресії, CR	Рівень декомпресійних спотворень, PSNR
512 x 512	4 x 4	4	4 : 1	26,0890408334672
512 x 512	8 x 8	5	13 : 1	24,0141453547684
512 x 512	8 x 8	4	16 : 1	23,3447304582785
512 x 512	8 x 8	3	21 : 1	22,5610068669955
512 x 512	16 x 16	20	18 : 1	39,7747328371311
512 x 512	16 x 16	10	26 : 1	30,1317094760289
512 x 512	16 x 16	9	28 : 1	28,7654800308728
512 x 512	16 x 16	8	32 : 1	27,5703466835148
512 x 512	16 x 16	4	64 : 1	23,4425350033892

Таблиця 4.

## Результати кодування зображення "Mandrill\_color"

Розмір зображення	Розмір блоку зображення	Кількість нейронів прихованого шару	Степінь компресії, CR	Рівень декомпресійних спотворень, PSNR
512 x 512	4 x 4	4	4 : 1	24,9350380414883
512 x 512	8 x 8	5	13 : 1	22,8036112681625
512 x 512	8 x 8	4	16 : 1	22,1897603681518
512 x 512	8 x 8	3	21 : 1	21,4841015184737
512 x 512	16 x 16	20	18 : 1	32,184541410029
512 x 512	16 x 16	10	26 : 1	25,9898208447403
512 x 512	16 x 16	9	28 : 1	25,302694456698
512 x 512	16 x 16	8	32 : 1	24,6290129970299
512 x 512	16 x 16	4	64 : 1	21,7546214229388

## Результати кодування зображення "Barbara\_gray"

Розмір зображення	Розмір блоку зображення	Кількість нейронів прихованого шару	Степінь компресії, CR	Рівень декомпресійних спотворень, PSNR
512 x 512	4 x 4	2	8 : 1	27,4474401274541
512 x 512	8 x 8	5	13 : 1	29,2700355390758
512 x 512	8 x 8	4	16 : 1	27,9789061665093
512 x 512	8 x 8	3	21 : 1	26,4580492642247
512 x 512	16 x 16	14	18 : 1	48,2942809520147
512 x 512	16 x 16	10	26 : 1	36,8653474189196
512 x 512	16 x 16	9	28 : 1	34,9473249251018
512 x 512	16 x 16	8	32 : 1	33,2203129525306
512 x 512	16 x 16	4	64 : 1	26,6743180245052

У табл. 3 наведено результати кодування чорно-білого тестового зображення "Mandrill". Порівнюючи отримані після декодування показники рівня декомпресійних спотворень із результатами застосування класичних нейромережних схем ущільнення зображень (осовський), можемо зробити висновки, що метод оснований на нейроподібних структурах МГП, переважає за відповідними показниками нейромережі традиційної архітектури на 5.21 %.

## Висновки

1. Розроблений метод компресії зображень, що забезпечує представлення скомпресованих даних у форматі чисел з фіксованою комою, вперше для нейромережних методів реалізує принципову можливість компресії даних без втрат або з регульованими втратами.

2. Розроблений метод дав змогу отримати показники якості компресії (беручи до уваги рівень декомпресійних спотворень), що переважають відповідні показники для нейромереж традиційної архітектури.

3. Подальше покращання якості компресії реалізується шляхом розбиття зображення на фрейми та регулювання довжини розрядної сітки; вказаний підхід забезпечується за рахунок високої швидкості навчання нейроподібних структур геометричних перетворень.

1. Verma B., Blumenstein M., Kulkarni S. "A new compression technique using an artificial neural network". Faculty of Information and Communication Technology, Griffith University, Australia, 2004.
2. Egmont-Ptersen M., D. de Ridder, Handels H., "Image processing with neural networks- a review", Pattern Recognition, 35 (2002) pp. 2279- 2301.
3. "Digital Imaging & Data compression; Image compression with neural networks".
4. Jiang J. Image compression with neural networks – A survey, Signal Processing: Image Communication 14 (1999) pp. 737 – 760.
5. Egmont-Ptersen M., D. de Ridder, Handels H. Image processing with neural networks- a review, Pattern Recognition , 35 (2002) pp. 2279-2301.
6. Ткаченко Р.О. Нова парадигма штучних нейронних мереж прямого поширення // Вісник Держ. ун-ту "Львівська політехніка": Комп'ютерна інженерія та інформаційні технології. – 1999. – №386. – С. 43–54.
7. Грицик В.В., Ткаченко Р.О. Нові підходи до навчання штучних нейромереж // Доповіді Національної академії наук України. – 2002. – № 11. – С. 59–65.
8. Ткаченко Р.О. Автореф. дис. ... д-ра техн. наук «Нейронні мережі прямого поширення з неітераційним навчанням». – Л., 2000.
9. Уоссермен Ф. Нейрокомпьютерная техника. – М.: Мир, 1992.
9. Осовский С.. Нейронные сети для обработки информации. – М.: Финансы и статистика, 2004