



О.Г. Руденко, М.С. Сныткин

Харьковский национальный университет радиоэлектроники,  
г. Харьков, Украина

## СЖАТИЕ ИЗОБРАЖЕНИЙ С ПОМОЩЬЮ НЕЙРОННОЙ СЕТИ FUZZY ART

Рассматривается JPEG-подобный алгоритм сжатия видеоинформации на основе нейронной сети Fuzzy ART, осуществляющей операцию векторного квантования. Результаты моделирования предлагаемого алгоритма в среде Matlab свидетельствуют о возможности применения данного алгоритма для сжатия изображения и целесообразности его использования при сжатии изображений, имеющих повторяющиеся участки.

АЛГОРИТМ СЖАТИЯ, ВЕКТОРНОЕ КВАНТОВАНИЕ, ДИСКРЕТНО-КОСИНУСНОЕ ПРЕОБРАЗОВАНИЕ, НЕЙРОННАЯ СЕТЬ FUZZY ART

## Введение

Увеличение объемов зрительной информации, хранимой, передаваемой и обрабатываемой в цифровом виде, обуславливает потребность в ее компактном представлении. В связи с этим возникает необходимость постоянного совершенствования алгоритмов и методов обработки информации.

Одним из перспективных представляется развитие подхода, в основе которого лежат искусственные нейронные сети (ИНС). При этом ИНС могут использоваться как при сжатии без потерь (например, в статистических методах кодирования для оценки вероятностей появления символов [1]), так и при реализации сжатия с потерями. В последнем случае целесообразно применение ИНС, осуществляющих векторное квантование.

В данной статье рассматривается применение ИНС на основе теории адаптивного резонанса – Fuzzy ART для развития JPEG-подобных алгоритмов сжатия видеоинформации.

## 1. Описание схемы сжатия

Общая схема сжатия видеоинформации состоит из таких основных шагов: дискретно-косинусное преобразование, векторное квантование, дифференциальное кодирование, энтропийное кодирование.

*Дискретно-косинусное преобразование* (ДКП) по сути представляет собой разложения участков изображения по амплитудам некоторых частот. Для сокращения объема хранимой информации используется тот факт, что в матрице частот многие коэффициенты либо близки, либо равны нулю [2].

*Векторное квантование*, или квантование коэффициентов частотной матрицы, представляет собой операцию, которая определяет степень сжатия и потери информации, увеличивает количество одинаковых и равных нулю элементов. При этом из пространства выбирается ограниченное число векторов (кодовых слов) для наиболее точного отображения распределения исходных векторов, полученных из представленного в виде блоков

изображения. Таким образом, большее количество кодовых слов будет расположено в той области пространства, где находится больше точек в исходном распределении (изображении) и наоборот. Затем каждый вектор исходного изображения заменяется ближайшим кодовым словом. На последнем этапе индекс кодового слова, содержащий меньшее количество бит, чем количество бит в исходном блоке, передается вместо самого слова.

*Дифференциальное кодирование.* Если предположить, что большинство частей изображения имеют плавные переходы, то дифференциальное кодирование, примененное после векторного квантования, приведет к «уменьшению» кодов в целом.

*Энтропийное кодирование* – кодирование словами (кодами) переменной длины, при которой длина кода символа имеет обратную зависимость от вероятности появления символа в передаваемом сообщении [3].

Кодирование длин серий является очень простой формой сжатия данных, в которой серии данных, т. е. последовательности, в которых одно и то же значение данных встречается во многих последовательных элементах данных, сохраняются в виде одиночного значения данных и их количества вместо оригинальной серии. Это сжатие часто используется для данных, содержащих много таких серий (например, простые графические изображения, такие как иконки и графические рисунки).

## 2. Сеть Fuzzy ART

Сеть Fuzzy ART является расширением сети ART1 путем применения теории нечетких множеств, что позволяет новой сети работать как с бинарными, так и с аналоговыми входными образами [4]. Для Fuzzy ART основные фазы классификации следующие.

*Предварительная обработка.* Все величины входного образа должны быть в интервале  $[0,1]$

$$i_k \in [0,1] \quad \forall k. \quad (1)$$

*Распознавание.* Восходящая сетевая активность, ведущая к предварительному выбору прототипа,

определяется с использованием нечеткой конъюнкции ( $\wedge$ ), по формулам:

$$\begin{aligned} x \wedge y &= \min\{x, y\} \\ X \wedge Y &= (x_1 \wedge y_1, \dots, x_m \wedge y_m), \end{aligned} \quad (2)$$

где  $Y$  – нечеткое подмножество  $X$ , если  $X \wedge Y = Y$ . Размер вектора ( $|X|$ ) определяется его нормой L1, т. е. суммой его элементов.

Активность  $t_j$ , каждого нейрона можно рассматривать как степень принадлежности прототипа  $W_j$  нечеткому подмножеству входного образа  $I$

$$t_j = \frac{|I \wedge W_j|}{\alpha + |W_j|}, \quad (3)$$

где  $\alpha = const$  – величина, играющая регуляризующую роль, т.е. предотвращающая возникновение переполнения при операции деления при  $|W_j| \rightarrow 0$ .

*Сравнение.* Сходство между входом  $I$  и победившим прототипом  $W_j$  определяется степенью принадлежности образа  $I$  нечеткому подмножеству  $W_j$ . Резонанс и адаптация происходят, если

$$\rho \leq \frac{|I \wedge W_j|}{|I|}. \quad (4)$$

*Адаптация.* Адаптация победившего прототипа  $W_j$  происходит путем изменения его компонентов по отношению к вектору  $I \wedge W_j$ :

$$W_j^{(new)} = \eta \cdot (I \wedge W_j^{(old)}) + (1 - \eta) \cdot W_j^{(old)}, \quad (5)$$

где  $\eta \in [0, 1]$  – показатель обучения, определяющий скорость сходимости прототипов к общему минимуму значений элементов всех входных образов, принадлежащих одному классу.

Сеть ART может работать в режиме классификации, если для предварительно обученной сети установить  $\eta = 0$ , что предотвратит модификацию прототипов новыми входными образами. Начальная инициализация прототипов выполняется постоянной величиной

$$w_{ij} \geq 1 \quad \forall i. \quad (6)$$

Таким образом обеспечивается поиск сначала среди фиксированных прототипов, а затем - среди остальных. Часто используемый метод ускорения обучения в сетях ART это установка коэффициента обучения  $\eta = 1$ , когда прежде неиспользованный прототип адаптируется к текущему входному вектору. Входной вектор  $I$  становится первым прототипом в новом классе, если другие ранее сформированные прототипы не подходят. Однако уже сформированные прототипы должны адаптироваться более медленно ( $\eta < 1$ ), чтобы предотвратить их искажение зашумленными входными образами.

*Дополнительное кодирование.* В сети Fuzzy ART существует проблема кластерного распространения, состоящая в том, что поскольку векторные

элементы прототипов после адаптации только уменьшаются, сеть стремится создавать больше прототипов, которые соответствуют входным образам с большими значениями входных величин, тогда как прототипы с малыми значениями могут никогда не быть доступны. Это устраняется путем нормализации, например, путем нормализации входных образов.

Обычно используется модифицированный вариант нормализации, называемый дополнительным кодированием, который преобразовывает все входные образы к одинаковой длине вектора. При этом оригинальный вектор  $A = (a_1, \dots, a_k)$  кодируется во входной образ  $I = (i_1, \dots, i_m)$  с добавлением своих дополнительных элементов к оригинальному вектору. Это удваивает длину всех входных образов и прототипов

$$I = (A, A^C) = (a_1, \dots, a_k, 1 - a_1, \dots, 1 - a_k) \quad a_i \in [0, 1] \quad \forall i.$$

Норма L1 векторов, закодированных этим методом и имеющих одинаковую длину, является величиной постоянной, независимой от величин элементов

$$\begin{aligned} |I| &= \sum_{i=1}^{2k} i_i = \sum_{i=1}^k a_i + \sum_{i=1}^k (1 - a_i) = \\ &= \sum_{i=1}^k a_i + k - \sum_{i=1}^k a_i = k = m / 2. \end{aligned}$$

Использование дополнительного кодирования упрощает выражение (4):

$$\rho \leq \frac{|I \wedge W_j|}{k}.$$

Начальная инициализация прототипов происходит в соответствии с (6). При работе в режиме быстрого обучения ( $\eta \rightarrow 1$ ) прототип  $W_j$  в Fuzzy ART является общим MIN-вектором для всех входных образов  $I_p$ ,  $p = (1, \dots, l)$ , входящих в один и тот же класс  $j$ , и может быть представлен в таком виде:

$$\begin{aligned} W_j &= (I_1, I_2, \dots, I_l) = \\ &= (\min\{i_{11}, \dots, i_{l1}\}, \dots, \min\{i_{1m}, \dots, i_{lm}\}). \end{aligned}$$

Использование дополнительного кодирования входных образов  $I_p = (A_p, A_p^C)$  приводит к прототипам, представляющим общий MIN- и MAX- вектор всех некодированных образов  $A_p$ :

$$\begin{aligned} W_j &= (\min\{A_1, \dots, A_l\}, \min\{A_1^C, \dots, A_l^C\}) = \\ &= (\min\{A_1, \dots, A_l\}, \max\{A_1, \dots, A_l\}^C). \end{aligned}$$

При малых значениях  $\eta$  сетевые прототипы сходятся более медленно к этим MIN и MAX границам.

### 3. Алгоритм сжатия

Алгоритм сжатия включает следующие шаги:

1. Условное разбиение исходного изображения на квадратные блоки размером  $n \times n$  пикселей, например,  $4 \times 4$  или  $8 \times 8$ .

2. Применение к каждому блоку дискретно-косинусного преобразования.

3. Представление каждого блока в виде вектора в 16- или 64-мерном пространстве соответственно.

4. Низкочастотная фильтрация (исключение из векторов коэффициентов, соответствующих высокочастотным составляющим).

5. Обучение нейронной сети ART.

6. Считывание индексов нейронов, соответствующих каждому входному вектору.

7. Формирование таблицы соответствия между индексом и усредненным вектором (кластерным центром).

8. Сжатие последовательности индексов с использованием кодирования длин серий и алгоритма Хаффмана.

Как следует из описания алгоритма, в нем, в отличие от JPEG, применена сеть ART (шаги 5-7), результаты работы которой используются для векторного квантования.

Для декодирования используются соответствующие операции, но в обратном порядке.

#### 4. Моделирование

При моделировании в среде Matlab процесса сжатия для оценки качества сжатия использовались показатели PSNR (peak signal-to-noise ratio) – отношение максимально возможного уровня сигнала к уровню искажающего его шума (поскольку большинство сигналов имеют очень широкий динамический диапазон, PSNR обычно представляют в логарифмическом масштабе), и MSE (mean squared error) – среднеквадратичная ошибка.

Для двух монохромных изображений  $I$  и  $K$  размерностью  $m \times n$  (где одно из изображений является зашумленным представлением второго) MSE вычисляется следующим образом:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \|I(i, j) - K(i, j)\|^2.$$

Для цветных изображений с тремя RGB компонентами MSE определяется как сумма всех квадратичных разностей, деленная на размер изображения и на 3.

Показатель PSNR определяется так:

$$PSNR = 10 \cdot \log_{10} \left( \frac{MAX_I^2}{MSE} \right) = 20 \cdot \log_{10} \left( \frac{MAX_I}{\sqrt{MSE}} \right),$$

где  $MAX_I$  представляет собой максимальное значение пикселя в изображении. В случае, если пиксели представлены 8-ми битовыми значениями,  $MAX_I = 255$ . В общем случае при использовании для представления  $B$  бит максимально возможное значение для  $MAX_I$  равно  $2^B - 1$ .

Обычно для алгоритмов сжатия PSNR находится в диапазоне 30-40 dB.

Для моделирования процесса сжатия были использованы блоки размером  $8 \times 8$  точек;

квантование осуществлялось при помощи сети Fuzzy ART. Для иллюстрации полученных результатов используется общепринятое изображение «Lena». В табл. 1 представлены результаты моделирования предложенного алгоритма, а также, для сравнения результатов сжатия с другими алгоритмами, представлены данные для алгоритмов JPEG и ART2 взятые из [5].

На рис. 1 и 2 приведены обобщенные графики зависимости PSNR и MSE от степени сжатия (CR) для всех алгоритмов.

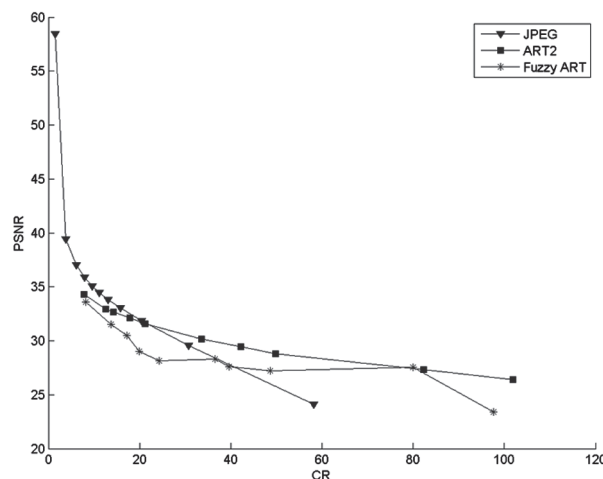


Рис. 1. Зависимости MSE от степени сжатия

Таблица 1

	Q/ρ	CR	PSNR	MSE
JPEG	0	58,215	24,110	252,389
	10	30,692	29,566	71,851
	20	20,403	31,848	42,490
	30	15,750	33,042	32,274
	40	13,078	33,814	27,023
	50	11,178	34,444	23,371
	60	9,611	35,059	20,284
	70	7,863	35,889	16,755
	80	6,073	37,046	12,837
	100	1,481	58,440	0,093
ART2	0,950	101,893	26,395	149,127
	0,960	82,319	27,356	119,519
	0,970	49,762	28,821	85,298
	0,975	42,235	29,467	73,513
	0,980	33,552	30,184	62,331
	0,984	21,209	31,590	45,090
	0,986	17,818	32,113	39,971
	0,988	14,296	32,691	34,990
	0,989	12,602	32,931	33,110
	0,990	7,853	34,330	23,994
Fuzzy ART	0,970	97,633	23,396	297,49
	0,980	79,983	27,525	114,97
	0,982	48,626	27,211	123,58
	0,984	39,585	27,613	112,66
	0,986	36,596	28,296	96,263
	0,988	24,274	28,166	99,186
	0,990	19,964	29,021	81,473
	0,992	17,191	30,475	58,29
	0,994	13,678	31,528	45,739
0,996	8,1339	33,596	28,409	

На рис. 3 показан пример сжатия изображения «Lena» с помощью сети Fuzzy ART при задании параметра сходства  $\rho = 0,986$  и показателя обучения  $\eta = 0,25$ . Степень сжатия в этом случае равна 36,6, а PSNR – 28,3 dB.

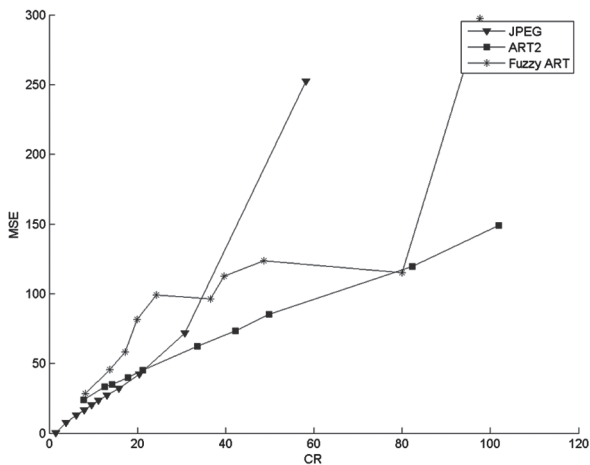


Рис. 2. Зависимости PSNR от степени сжатия



Рис. 3. Пример сжатия изображения

### Выводы

В данной статье представлен новый алгоритм сжатия изображений на основе нейросетевого подхода. Основной его отличительной особенностью является то, что на этапе векторного квантования для создания соответствия между входным пространством и выходным пространством, состоящим из элементов кодовой книги - кодовых слов, или нейронов, была применена ИНС Fuzzy ART. Результаты моделирования данного алгоритма свидетельствуют о возможности применения данного вида ИНС для сжатия изображений.

Рассмотренный алгоритм сжатия более эффективен, чем алгоритм, основанный на использовании карт Кохонена и описанный в [6], и показал аналогичные результаты с алгоритмами, описанными в [5]. Дальнейшее направление исследований видится в применении различных модификаций

алгоритмов работы сети Fuzzy ART, таких как стратегия голосования, алгоритм упорядочивания, модифицированный алгоритм обучения [7-9] для дополнительного увеличения степени сжатия изображений.

**Список литературы:** 1. Schmidhuber J. Sequential Neural text Compression / J. Schmidhuber, S. Heil // IEEE Trans. on Neural Networks. – 1996. – V. 7. – N.1 – P.142-146. 2. Ahmed N. Discrete cosine transform / N. Ahmed, T. Natarajan and K. R. Rao // IEEE Trans. Comput. – 1974. – C-23 – P. 90-93. 3. Lei Sh.-M. An entropy coding system for digital HDTV applications / Sh.-M. Lei, M.-T. Sun // IEEE Trans. Circuits Syst. Video Technol. – 1991. – 1 – P. 147-155. 4. Carpenter G. A. Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system / G. A. Carpenter, S. Grossberg, D. B. Rosen // Neural Networks. – 1991. – Vol. 4. – P. 759-771. 5. Руденко, О. Г. Сжатие изображений на основе нейронной сети ART [Текст] / О. Г. Руденко, М. С. Сныткін // Кибернетика и системный анализ. – 2008. – №6. – С.10-16. 6. Amerijckx Ch. Image Compression by Self-Organized Kohonen Map / Ch. Amerijckx, M. Verleysen, P. Thissen, J.-D. Legat // IEEE Trans. on Neural Net. – 1998. – 9 – P. 503-507. 7. Dagher I. An ordering algorithm for pattern presentation in Fuzzy ARTMAP that tends to improve generalization performance / I. Dagher, M. Georgiopoulos, G. L. Heileman, G. Bebis // IEEE Transactions on Neural Networks. – 1999. – Vol. 10, №4. – P. 768-778. 8. Meneganti M. Fuzzy neural networks for classification and detection of anomalies / M. Meneganti, F. S. Saviello, R. Tagliaferri // IEEE Transactions on Neural Networks. – 1998. – Vol. 9, №5. P. 848-861. 9. Rizzi A. A recursive algorithm for fuzzy min-max networks / A. Rizzi, M. Pannela, F. M. Fratalle Mascioli, G. Martinelli // Proceedings of International Joint Conference on Neural Networks (IJCNN'2000). – 2000. – Vol. 6. – P. 541-546.

Поступила в редколлегию 23.11.2011

УДК 519.71

**Стиснення зображень за допомогою нейронної мережі Fuzzy ART** / О. Г. Руденко, М. С. Сныткін // Біоніка інтелекту: наук.-техн. журнал. – 2012. – № 1 (78). – С. 45-48.

В статті розглядається алгоритм стиснення зображень на основі нейронної мережі Fuzzy ART, яка виконує операцію векторного квантування. Наведено опис алгоритму роботи і навчання нейронної мережі та алгоритму стиснення зображень. Також наводяться результати моделювання, які свідчать про ефективність застосування запропонованого алгоритму для стиснення зображень.

Табл. 1. Іл. 3. Бібліогр.: 9 найм.

UDC 519.71

**Image compression by neural network Fuzzy ART** / O. Rudenko, M. Snytkin // Bionics of Intelligense: Sci. Mag. – 2012. – № 1 (78). – P. 45-48.

In this paper an algorithm of image compression is considered. The algorithm is based on Fuzzy ART neural network. The network is used for vector quantization. The network's structure, its training algorithm and algorithm of image compression are described. The simulation results that show efficiency of the proposed algorithm of image compression are presented.

Tab. 1. Fig. 3. Ref.: 9 items.