

СЖАТИЕ ИЗОБРАЖЕНИЙ НА ОСНОВЕ НЕЙРОННОЙ СЕТИ ART

Ключевые слова: *нейронная сеть, алгоритм сжатия, векторное квантование, обучение.*

ВВЕДЕНИЕ

Непрерывное возрастание объемов перерабатываемой информации, в частности зрительной, обуславливает необходимость ее компактного представления. Применение с этой целью традиционных методов, использующих префиксное либо арифметическое кодирование и обеспечивающих сжатие без потерь, требует знания статистики (частот) появления элементов. Реализация весьма эффективных статистических методов кодирования, использующих оценки условных вероятностей появления символов в различных контекстах, требует значительных вычислительных ресурсов [1]. Проблема применения комбинаторного подхода, основанного на алгоритмах А. Лемпела и Я. Зива, используемых для определения повторяющихся подстрок (надстрок), и рассматривающего сообщение как цепочку символов, обусловлена тем, что изображение имеет не только ширину, но и высоту, причем вероятность появления вертикальных повторяющихся строк равна вероятности появления горизонтальных строк.

В то же время наличие в мультимедийной информации операции дискретизации изображения и звука позволяет ставить и решать задачу эффективного сжатия информации с потерями. Во многих приложениях, например при представлении неподвижных изображений, такое сжатие более чем оправданно. Наиболее популярный в настоящее время стандарт JPEG включает оба типа сжатия, однако зачастую не удовлетворяет возрастающим требованиям функциональности.

В этих условиях перспективным 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 [2]), так и при реализации сжатия с потерями, например, в стандарте JPEG 2000, основанном на вейвлет-преобразовании. В последнем случае целесообразно применение ИНС, осуществляющих векторное квантование.

Наиболее значительные результаты в этом направлении получены на основе самоорганизующихся карт Кохонена [3] и сетей встречного распространения, также содержащих слой нейронов Кохонена [4, 5]. Самоорганизация в этих сетях представляет собой процесс кластеризации образов, осуществляемой аналогично методу главных компонент. Следует отметить, что и для реализации самого метода главных компонент могут успешно применяться нейросетевые технологии, в частности нейронная сеть Т. Сэнгера [6], состоящая из нейронов Э. Оя [7, 8].

В данной статье рассматривается применение ИНС, построенной на основе предложенной и развитой С. Гроссбергом [9] теории адаптивного резонанса (ART), для развития JPEG-подобных алгоритмов сжатия видеoinформации.

ИНС ART относит входное изображение к одному из известных классов, если оно в достаточной степени подобно или резонирует с прототипом этого класса. Если найденный прототип с определенной точностью, задаваемой специальным параметром сходства $\rho \in (0, 1]$, соответствует входному изображению, то он модифицируется, чтобы стать более похожим на предъявленное изображение. Когда входное изображение недостаточно подобно ни одному из имеющихся прототипов, на его

основе создается новый класс. Это возможно благодаря тому, что сеть имеет большое число избыточных или нераспределенных элементов, которые не используются до тех пор, пока в этом нет необходимости (если нет нераспределенных нейронов, то входное изображение не вызывает реакции сети), благодаря этому новые образы могут создавать новые классы, не искажая запомненную информацию.

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) представляет собой по сути разложение 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 [10].

Векторное квантование. 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 битов по сравнению с исходным блоком, 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 [11]. Кодирование длин серий — простая форма 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: иконки и графические рисунки).

СЕТЬ ART

Функционирование данной сети, как и всякой ИНС, начинается с ее обучения. Полученный после обучения набор нейронов должен максимально точно отображать векторы из входного пространства [12].

В зависимости от вида входных переменных и способа их обработки применяют различные модели сетей ART, среди которых наибольшее распространение получили сети ART 2 (рис. 1).

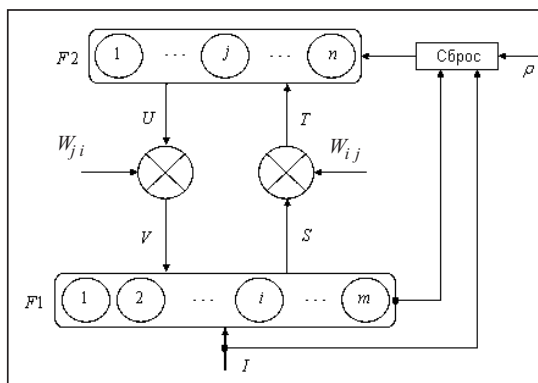


Рис. 1. Схема упрощенной архитектуры сети ART 2

Начальная обработка в любой сети ART выполняется в модуле, представляющем собой обучаемую конкурентную сеть. Входные m нейронов слоя $F1$ хранят входной образ $I = (i_1, i_2, \dots, i_m)$. На каждый нейрон выходного слоя $F2$ поступает восходящая сетевая активность t_j , формируемая из всех выходов $S = I$ слоя $F1$.

Элементы вектора $T = (t_1, t_2, \dots, t_n)$, вычисляемые как $t_j = \sum_{i=1}^m w_{ij} \cdot i_i$, можно рас-

сматривать как результат сравнений входного образа I с прототипами $W_1 = (w_{11}, \dots, w_{1m}), \dots, W_n = (w_{n1}, \dots, w_{nm})$ — весами синаптических связей между слоями $F1$ и $F2$. Выход только одного нейрона J слоя $F2$, получившего наибольшую восходящую сетевую активность t_j , устанавливается в единицу, на выходах же остальных нейронов остаются нули.

После того как определен нейрон-победитель J слоя $F2$, соответствующий прототип $W_J = (w_{1J}, \dots, w_{mJ})$ адаптируется к входному образу I согласно формуле

$$W_J^{(\text{new})} = \eta \cdot I + (1 - \eta) \cdot W_J^{(\text{old})},$$

где $\eta \in [0, 1]$ — показатель обучения.

Синаптические нисходящие веса W_{ji} обратных связей, за исключением возможного коэффициента масштабирования, идентичны восходящим весам W_{ij} , а нисходящая сетевая активность V определяется по аналогии с восходящей:

$$v_i = \sum_{j=1}^n u_j \cdot w_{ji}.$$

Поскольку все выходы слоя $F2$, кроме одного — u_j , равны нулю, входной слой $F1$ получает прототип W_J , который представляет текущий класс J , победивший в соревнованиях. Далее выполняется самая сложная в сетях ART часть обработки данных — сравнение прототипа W_J с входным образом I . Задаваемая исследователем чувствительность ρ определяет требуемое минимальное сходство между входным образом и прототипом соответствующего класса. Если степень совпадения меньше ρ , текущий нейрон-победитель слоя $F2$ исключается из соревнований сигналом сброса. Сигнал сброса устанавливает активный нейрон J слоя $F2$ в нуль и таким образом дает возможность другому нейрону победить в соревнованиях, тем самым на выходе слоя $F2$ получаем восходящую сетевую активность t_j еще не сброшенных выходных нейронов. Как только будет обнаружен прототип, у которого степень сходства с входным образом I , по крайней мере, такая же, как параметр сходства ρ , сигнал сброса не возникнет и сеть достигнет резонанса. Позиция последнего нейрона-победителя слоя $F2$ указывает класс входного образа I , после чего происходит адаптация соответствующего прототипа.

АЛГОРИТМЫ СЖАТИЯ

В предлагаемых ниже алгоритмах сжатия вместо использования простой схемы дифференциального кодирования (предиктор нулевого порядка), в которой каждый новый код вычитается из кода, соответствующего предыдущему закодированному блоку изображения, т.е. блоку, находящемуся левее от текущего блока, используется следующий принцип (предиктор первого порядка). Если предположить, что переходы в изображении гладкие, то направление, в котором разность между кодами двух уже закодированных блоков минимальна, будет такое же, как и направление, в котором разность минимальна для нового кодируемого блока. Другими словами (рис. 2), предполагается, что минимальная разность между кодами блоков i и b , i и d , i и f , i и h будет соответственно в таком же направлении ($D1, D2, D3$ и $D4$), как и разность между кодами закодированных блоков b и a , d и c , f и e , h и g .

Зачастую, если изображение имеет гладкие переходы (что является обычным для большинства изображений, за исключением областей с резкими переходами, где дифференциальная схема не приносит каких-либо преимуществ), выбор между четырьмя возможными направлениями дает меньшую разность по сравнению с общепринятой схемой. Кроме того, выбранное направление не надо кодировать, поскольку коды блоков $a-h$ были уже переданы до кодирования блока i , а направление минимальной разности между кодами можно вычислить из существующих блоков и поэтому не требует дополнительной передачи данных.

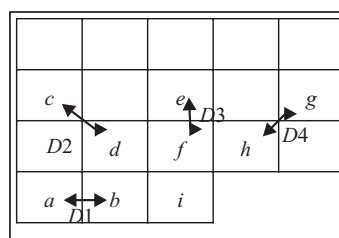


Рис. 2. Схема определения лучшего направления для кодирования

Алгоритм 1. Наиболее сложный и включает следующие шаги:

- 1) условное разбиение исходного изображения на квадратные блоки размером $n \times n$ пикселей, например, 4×4 или 8×8 ;
- 2) применение к каждому блоку ДКП;
- 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.

Алгоритм 2. Включает те же шаги, что и алгоритм 1, за исключением шагов 2 и 4. Таким образом, в данном алгоритме сеть выполняет свою основную функцию — осуществляет векторное квантование без вычисления ДКП и низкочастотной фильтрации.

Алгоритм 3. Данный алгоритм основан на алгоритме 2, но для повышения качества изображения добавлен еще один цикл сжатия, в котором происходит кодирование ошибки изображения, т.е. из разности между исходным изображением и изображением, полученным после первого цикла кодирования. При декодировании второе изображение суммируется с первым, таким образом происходит коррекция ошибки кодирования.

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 является зашумленным представлением второго) MSE вычисляется следующим образом:

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

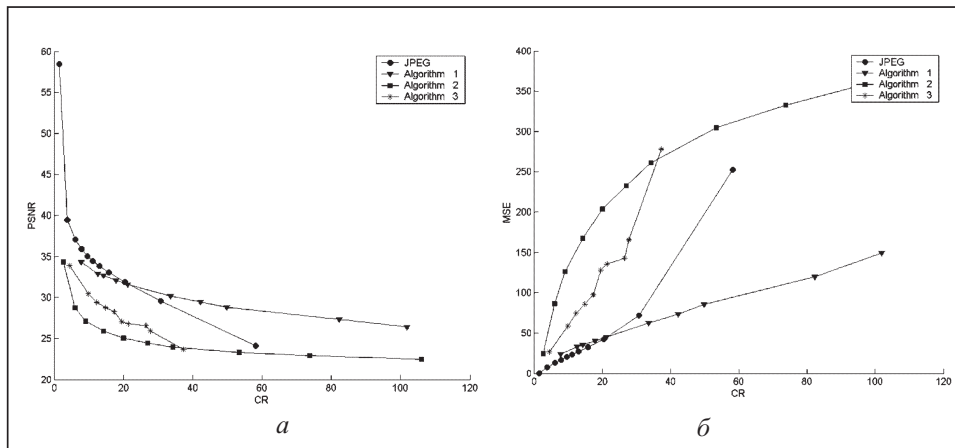


Рис. 3. Кривая зависимости PSNR (а) и MSE (б) от степени сжатия

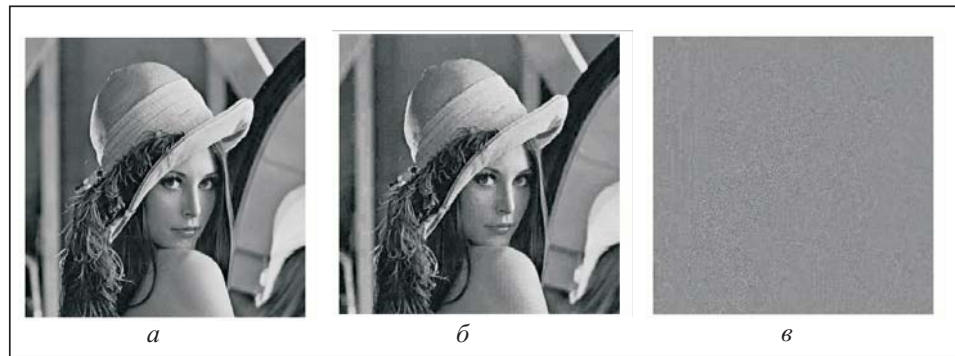


Рис. 4. Изображение: оригинальное (а); сжатое (б); поверхность ошибки (в)

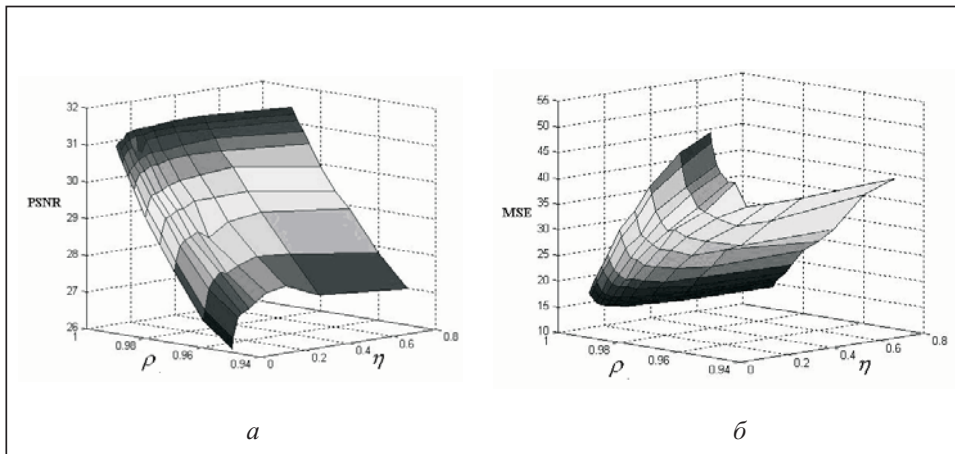


Рис. 5. Зависимости PSNR (а) и MSE (б)

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:

$$\text{PSNR} = 10 \cdot \log_{10} \left(\frac{\text{MAX}_I^2}{\text{MSE}} \right) = 20 \cdot \log_{10} \left(\frac{\text{MAX}_I}{\sqrt{\text{MSE}}} \right).$$

Здесь MAX_I — максимальное значение пиксела в изображении. 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 использовались блоки размером 8×8 точек; квантование осуществлялось с помощью сети ART 2A sub E. Предложенные алгоритмы сжатия показали схожие результаты для различных изображений, для иллюстрации полученных результатов ниже используется общепринятое изображение «Lena».

В табл. 1 представлены результаты моделирования алгоритмов 1–3 и алгоритма JPEG.

На рис. 3 приведены обобщенные графики зависимости PSNR и MSE от степени сжатия (CR) для всех алгоритмов.

На рис. 4 показан пример сжатия изображения «Lena» с помощью сети ART 2A–E при задании параметра сходства $\rho = 0,980$ и показателя обучения $\eta = 0,25$. Степень сжатия в этом случае равна 33,6, а PNSR = 30,2 dB.

Рис. 5 отражает зависимости PSNR и MSE от ρ и η . Следует отметить, что характер этих зависимостей практически одинаков для всех рассматриваемых сетей ART. Из рисунка видно, что PSNR прямопропорционально зависит от ρ , т.е. этот параметр непосредственно влияет на качество сжатого изображения. Как показало моделирование, при $\eta \approx 0,2$ эффективность обучения ИНС достигает своего максимума, что соответствует оптимальному соотношению компонент обучающего вектора и вектора прототипа класса на этапе адаптации.

ЗАКЛЮЧЕНИЕ

В данной статье предложены новые алгоритмы сжатия на основе нейросетевого подхода. Для создания соответствия между входным и выходным пространствами, состоящими из элементов кодовой книги — кодовых слов, или нейронов, была применена ИНС ART. Результаты моделирования данных алгоритмов свидетельствуют о возможности применения данного вида ИНС для сжатия изображений.

Таблица 1. Результаты моделирования алгоритмов

Алгоритм	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
	90	3.823	39.440	7.398
	100	1.481	58.440	0.093
Алгоритм 1	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
Алгоритм 2	0.700	105.960	22.466	368.512
	0.740	73.760	22.908	332.855
	0.780	53.477	23.292	304.720
	0.820	34.231	23.964	261.018
	0.840	27.031	24.459	232.901
	0.860	20.008	25.040	203.745
	0.880	14.181	25.895	167.326
	0.900	9.053	27.122	126.135
	0.920	5.996	28.754	86.642
	0.960	2.658	34.304	24.134
Алгоритм 3	0.600	37.364	23.690	278.008
	0.640	27.817	25.943	165.477
	0.680	26.533	26.590	142.595
	0.720	21.292	26.804	135.722
	0.740	19.534	27.065	127.811
	0.760	17.287	28.257	97.145
	0.780	14.911	28.784	86.028
	0.800	12.368	29.408	74.525
	0.820	9.858	30.464	58.430
	0.860	4.399	33.893	26.530

Рассмотренные алгоритмы сжатия более эффективны, чем алгоритм, основанный на использовании карт Кохонена и описанный в [14], для которого при степени сжатия 25,22, PSNR составляет 24,7 dB. Кроме того, предложенные алгоритмы свободны от недостатка, связанного с необходимостью предварительного выбора количества кодовых слов, определяющих размер карты Кохонена.

Следует также отметить, что использование в алгоритмах дополнительного кодирования разностного изображения, т.е. ошибки начального кодирования, позволяет повысить качество получаемого изображения.

Применение разработанных алгоритмов наиболее эффективно при сжатии изображений, имеющих повторяющиеся участки.

СПИСОК ЛИТЕРАТУРЫ

1. Long P.M., Nastev A.I., Vitter J.S. Text compression via alphabet representation // Proc. IEEE Data Compres. Conf. — Snowbird, Utah, USA. 1997. — P.161–170
2. Schmidhuber J., Heil S. Sequential neural text compression // IEEE Trans. on Neural Networks. — 1996. — 7, N 1. — P. 142–146.
3. Kohonen T. Self-organizing maps. — Berlin: Springer-Verlag, 1995. — 362 p.
4. Heht-Nielsen R. Application counterpropagation networks // Neural Networks. — 1988 — 1. — P. 131–139
5. Heht-Nielsen R. Counterpropagation networks // Appl. optics. — 1987. — 26. — P. 4979–4984
6. Sanger T. Optimal unsupervised learning in a single-layer linear feedforward neural network // Neural Networks. — 1983. — 2. — P. 459–473.
7. Oja E. A simplified neuron model as a principal component analyzer // J. Math Biology. — 1982. — 15. — P. 267–273
8. Бодянский Е. В., Руденко О. Г. Искусственные нейронные сети: архитектура, обучение, применение. — Харьков: ТЕЛЕТЕХ, 2004. — 372 с.
9. Grossberg S. Competitive learning: from interactive action to adaptive resonance // Cognitive Science. — 1987. — 11. — P.23–63
10. Ahmed N., Natarajan T., Rao K.R. Discrete cosine transform // IEEE Trans. Comput. — 1974. — C-23. — P. 90–93.
11. Lei Sh.-M., Sun M.-T. An entropy coding system for digital HDTV applications // IEEE Trans. Circuits Syst. Video Technol. — 1991. — 1. — P. 147–155.
12. Бодянский Е.В., Кулишова Н.Е., Руденко О.Г. Обобщенный алгоритм обучения формального нейрона // Кибернетика и системный анализ. — 2002. — № 5. — С. 176–179.
13. Руденко О.Г., Бодянский Е.В. Искусственные нейронные сети — Харьков: Компания СМИТ, 2005. — 408 с.
14. Amerijckx Ch., Verleysen M., Thissen P., Legat J.-D. Image compression by self-organized Kohonen map // IEEE Trans. on Neural Net. — 1998. — 9. — P. 503–507.

Поступила 11.02.2008