УДК 681.32.05

О. М. КАРПОВА, А. И. ПИНЧУК, М. А. СТАРКОВ

(Новосибирск)

К ВОПРОСУ КОДИРОВАНИЯ ИЗОБРАЖЕНИЙ

С развитием цифрового телевидения задача сжатия цифровых телевизионных изображений привлекает к себе все большее внимание специалистов как в нашей стране [1], так и за рубежом [2]. Наиболее эффективными из предложенных в этих работах способов кодирования являются методы сжатия с применением ортогональных преобразований. Подробное исследование этих методов приведено в работе [3], где указывается плотность кода 2 бит/элемент при приемлемом качестве восстановленного изображения. При этом авторы считают достижимой плотность 1 бит/элемент. Следует заметить, что в этих методах оценка качества принятого изображения дается по среднеквадратическому критерию, который не позволяет выявить неизбежного размывания границ изображения. Кроме того, разбиение изображения на блоки приводит к появлению «шума фрагментарности» [4]. Для решения казалось бы противоречивой задачи сжатия изображения с одновременным улучшением его качества был разработан способ передачи, адаптивный к контурам изображения [4] и хорошо согласованный с особенностями зрительного восприятия. С помощью этого метода достигнута плотность кода 2,8 бит/элемент, причем принятое изображение визуально воспринималось лучше исходного. Авторами настоящей работы проведено исследование статистических свойств изображений и разработан способ математического описания изображений [5]. На основании предложенной модели реализован алгоритм кодирования, который позволил снизить плотность кода до 0,5 бит/элемент, причем состояние контурной части принятого изображения было лучше, чем исходного. К недостатку этого алгоритма следует отнести информационные потери, которые приводят к пропуску мелких деталей на изображении, разрыву линий. Однако вероятность пропуска мелких деталей можно заранее предсказать и ограничить, исходя из гарантированной полуширины автокорреляционной функции передаваемых изображений [6].

В настоящей работе рассматривается алгоритм кодирования, который дает возможность заранее ограничить вероятность пропуска градации и тем самым задать точность передачи. Получаемая при этом скорость передачи зависит от сложности или информационной емкости передаваемого изображения.

В работе [6] был описан следующий способ передачи изображений. Множество точек изображения разбивается на девять непересекающихся подмножеств $\Lambda_0, \Lambda_1, \ldots, \Lambda_8$. Множество Λ_0 есть объединение узлов квадратной сетки $(16j, 16k)$. Точки Λ_1 находятся в центрах квадратов с точками множества Λ_0 в углах и, таким образом, имеют координаты $(16j+8, 16k+8)$. Заметим, что $\Lambda_0 \cup \Lambda_1$ — также квадратная решетка, но с диагоналями, параллельными координатным осям. В центрах квадратов множества $\Lambda_0 \cup \Lambda_1$ располагаются точки множества Λ_2 и т. д. Объединение всех Λ_i $\bigcup\limits_{i=0}^{8} \Lambda_i$ включает в себя все изображение. Передача изображения осуществляется за девять актов, в каждом из которых последовательно пересылаются коды значений изображения в точках Λ_0, затем Λ_1 и т. д. Значения изображения в точках множества Λ_0 передаются в исходном виде (для определенности будем считать, что изображение представлено байтами). Для кодирования значений в точках Λ_i используем четверку значений в ближайших точках из множества $\bigcup\limits_{k=0}^{i-1} \Lambda_k$ и заметим, что значения указанной четверки на приемной стороне уже известны.

Теперь опишем процедуру кодирования, предварительно условившись, что два числа a и b равны, если $a \in [b-\delta,\ b+\delta]$. При кодировании точек, близких к границе матрицы, потребуются значения на точках, выходящих за размеры матрицы. В этих случаях будем считать, что матрица каким-либо образом продолжена, например, нулями. При кодировании значения в точке множества Λ_i выпишем четыре числа, представляющих значения изображения на четверке ближайших точек множества $\bigcup\limits_{k=0}^{i-1} \Lambda_i$, и вычеркнем из полученной записи повторяющиеся числа. При этом в записи может остаться m чисел $(1 \leqslant m \leqslant 4)$. Сравним значение в кодируемой точке с числами, оставшимися в записи. В этом случае могут возникнуть два варианта. Если кодируемое значение не совпадает ни с одним числом записи, код оставим без изменения (байтовая кодировка). Если кодируемое значение равно одному из чисел записи, то присвоим ему номер числа записи, с которой он совпал. При этом, если $2 < m \leqslant 4$, для записи номера отведем 2 бита, если $m=2$, — 1 бит, и, если $m=1$,— нуль битов, т. е. просто пустой код. Три значения реализуются на опорной четверке значительно чаще, чем четыре, причем вероятность совпадения кодируемого значения с тем, что реализовалось дважды, вдвое больше, чем с другими. Это дает возможность дополнительного сокращения кода, например, таким способом: если кодируемое значение совпадает с тем, что реализовалось дважды, ему присваивается код 1, если с каким-либо из оставшихся,— код 01 или 00 (предварительно следует условиться об отношении порядка между ними). Заметим, что все эти ситуации однозначно распознаются на приемной стороне, но байтовая кодировка не может быть распознана. Из r подряд кодируемых значений организуем группу первого типа:

$$0 n_1 n_2 n_3 \ldots n_r, \tag{1}$$

если в группе не встречается байтовая кодировка. Здесь одним битом перед группой записан нуль — признак группы первого типа, а n_i — код i-го значения в группе — может быть представлен 0,1 или 2 битами. В вырожденном случае длина группы равна одному биту, т. е. состоит только из признака типа группы. Если в группе имеется хотя бы одна байтовая кодировка, код группы будет таким:

$$1 0 n_1 0 n_2 \ldots 0 n_{j-1} 1 \text{Б}_j \ldots, \tag{2}$$

где первый бит, равный единице, есть признак группы второго типа, а перед кодом значения в точке стоит 0, если оно кодируется номером, и 1, если кодируется байтом; Б_j — байтовый код значения. Нетрудно заметить, что код всего изображения однозначно декодируется, если известна длина группы r. Теперь рассчитаем длину группы, минимизируя среднюю плотность кодирования. Обозначим через α_i вероятность совпадения кодируемого значения с одним из значений в ближайшей четверке точек множества $\bigcup\limits_{k=0}^{i-1} \Lambda_k$ в i-м акте кодирования. В [6] было показано, что α_i не зависит от реализации в четверке точек. При условии биномиального распределения байтовых кодов по группам вероятность реализации группы первого типа

$$P_1 = \alpha_i^r. \tag{3}$$

Вероятность реализации группы второго типа, содержащей в себе j байтовых кодов, определится выражением

$$P_{2j} = C_r^j \alpha_i^{r-j} (1-\alpha_i)^j. \tag{4}$$

Длина групп представляется в виде

$$L_1 = r\bar{n} + 1, \quad L_{2j} = (r-j)\bar{n} + 8j + r + 1, \tag{5}$$

где $\bar{n}$ — средняя длина кода точки при кодировании номером. Тогда средняя длина группы

$$L_c = P_1 L_1 + \sum_{j=1}^{r} P_{2j} L_{2j}. \tag{6}$$

Подставляя в (6) выражения (3)—(5), после несложных преобразований будем иметь

$$L_{ci} = r\bar{n}_i + r\left(1-\alpha_i^r\right) + (8-\bar{n}_i)(1-\alpha_i)\,r + 1. \tag{7}$$

Разделив это выражение на r, определим среднюю длину кодируемого элемента, или плотность кодирования, на i-м акте передачи:

$$\bar{l} = \bar{n}\alpha + 8(1-\alpha) + (1-\alpha^r) + 1/r. \tag{8}$$

Дифференцируя (8) по r и приравнивая производную нулю, получим

$$\alpha^r \ln\alpha + 1/r^2 = 0. \tag{9}$$

В табл. 1 приведены решения уравнения (9) для различных значений r, которые дают минимум выражению (8). Понятно, что α может принимать произвольные значения, отличные от указанных в табл. 1. Поэтому при получении промежуточных значений следует сделать выбор r либо в большую, либо в меньшую сторону. В четвертой колонке таблицы содержатся значения двух последних членов выражения (8) при выборе r на единицу меньше оптимального значения, в пятой колонке — на единицу больше оптимального значения. Эти результаты позволяют сделать вывод, что при получении промежуточных значений α целесообразнее сделать выбор r в сторону увеличения. Для $r = 1 \vee 2$ уравнение (9) не имеет действительных корней. Так как при любых значениях α выражение $(1-\alpha^r) + 1/r$ убывает с ростом r, то для тех из них, для которых не существует оптимального значения r, длина группы должна выбираться по возможности большей. Поскольку же вероятность реализации группы типа (1) при больших r близка к нулю и реализуются только группы типа (2), от признака типа группы можно отказаться, и тогда код акта передачи будет иметь следующий вид:

$$0n_1 0n_2 \ldots 0n_{j-1} 1\text{Б}_j \ldots, \tag{10}$$

и средняя скорость кодирования описывается выражением

$$\bar{l} = \bar{n}\alpha + 8(1-\alpha) + 1. \tag{11}$$

Теперь найдем, при каких значениях α следует переходить к кодам типа (10). Для этого сравним средние скорости кодирования, определяемые выражениями (8) и (11), при $r = 3$. Они будут одинаковыми при условии $\alpha^3 - 1/3 = 0$, или $\alpha = \sqrt[3]{1/3}$. Следовательно, при $\alpha < \sqrt[3]{1/3}$ от группирования следует отказаться и применять коды типа (10).

Очевидно, что $\bar{l}$, определенное из выражения (11), дает наибольшие значения для изображения, представляющего собой некоррелированный шум с равномерным распределением градаций яркости. Тот же результат получим, если расстояние между точками опорной четверки достаточно велико. В этом случае можно пренебречь вероятностью совпадения значений в точках опорной четверки между собой, и тогда $\bar{n} = 2$. Вероятность совпадения кодируемого значения с одним из значений опорной четверки будет равна $4/256 = 1/64$, а максимальное значение $\bar{l}_{\max} = 2/64 + 8(1 - 1/64) +$

Таблица 1

r	$\alpha_{\text{опт}}$	$1-\alpha^r+1/r$	$1-\alpha^{r+1}+1/(r+1)$	$1-\alpha^{r-1}+1/(r-1)$
3	0,814	0,794	0,811	0,838
4	0,915	0,549	0,560	0,567
5	0,949	0,448	0,463	0,466
6	0,967	0,349	0,352	0,356
7	0,976	0,299	0,302	0,305
8	0,982	0,260	0,262	0,265
9	0,987	0,222	0,224	0,226
10	0,989	0,205	0,206	0,207

Таблица 2

r	$\alpha_{опт}$	$m=1$		$m=2$		$m=3$		$m=4$	
		$\bar{l}$	H	$\bar{l}$	H	$\bar{l}$	H	$\bar{l}$	H
	0,167	7,667	7,310	7,834	7,456	7,918	7,551	8,000	7,630
	0,693	3,456	3,344	4,149	3,970	4,496	4,380	4,842	4,724
3	0,814	2,282	2,180	3,096	2,916	3,503	3,399	3,910	3,805
4	0,915	1,229	1,099	2,144	1,927	2,602	2,471	3,059	2,927
5	0,949	0,838	0,698	1,787	1,558	2,262	2,122	2,736	2,595
6	0,967	0,613	0,472	1,580	1,348	2,064	1,923	2,547	2,406
7	0,976	0,491	0,354	1,467	1,238	2,012	1,818	2,443	2,306
8	0,982	0,404	0,273	1,386	1,162	1,877	1,746	2,368	2,237
9	0,987	0,326	0,203	1,313	1,097	1,806	1,684	2,300	2,177
10	0,989	0,293	0,185	1,282	1,070	1,779	1,660	2,271	2,153

$+1=8{,}75$ бит. Таким образом, в самом неблагоприятном случае средняя плотность кодирования не превышает 8,75 бит/элемент, т. е. относительное удлинение кода не превосходит 10%. Представляется интересным рассмотреть такое значение α, при котором $\bar{l}=8$. Так же, как и для определения $\bar{l}_{\max}$, будем считать, что значения в опорной четверке попарно не совпадают друг с другом и $\bar{n}=2$; тогда, подставляя $\bar{n}=2$ и $\bar{l}=8$ в (11), находим $\alpha=1/6$.

Итак, от передающей стороны требуется выбор типа кода и длины группы для каждого акта передачи. Эти данные должны быть переданы по каналу связи либо заранее обусловлены при передаче определенных классов изображений. Выбор этих величин в зависимости от α может быть сделан по табл. 1.

Известно, что при оптимальном кодировании средняя скорость кодирования совпадает с энтропией. Для оценки эффективности работы вышеописанного кодера в табл. 2 приведены значения энтропии и средней длины кода для различных значений α при всех комбинациях опорных четверок. Способ вычисления энтропии описан в [6]. Из табл. 2 видно, что данный алгоритм кодирования дает достаточно близкий к оптимальному код при значениях $\alpha<0{,}95$. При больших значениях α, поскольку повышается частота реализации одного или двух значений на опорной четверке, уклонение средней длины кода от оптимальной значительно.

Рассмотрим меры, позволяющие повысить эффективность кодирования. В реализованной программе кодирования применяется два типа кодеров. Первый тип подробно описан выше. Второй тип кодера отличается от первого тем, что в нем отсутствует байтовое кодирование. В тех случаях, когда кодируемое значение не совпадает ни с одним из значений опорной четверки, оно будет заменено тем опорным, которое дает с ним наименьшую по модулю разность. Таким образом, в канал связи всегда будет отправлен только номер этого значения. При этом следует иметь в виду, что разность между кодируемым и переданным значениями может быть значительной. Второй тип кодера, как правило, применяется на седьмом и восьмом актах передачи. Оценим информационные потери на последнем акте. Решая для этого уравнение

$$\Delta P^r_{j,k}-\lambda\left(P^r_{j,k}-\hbar\right)=0, \qquad (j,k)\in\bar{\Omega};$$

$$P^r_{j,k}=\begin{cases}1, & \omega_{j,k}=r\\ 0, & \omega_{j,k}\neq r\end{cases}, \qquad (j,k)\in\Omega$$

относительно $P^r_{j,k}$, получим

$$P^r_{j,k}=\left(P^r_{j-1,k}+P^r_{j+1,k}+P^r_{j,k-1}+P^r_{j,k+1}+\lambda\hbar\right)/(4+\lambda). \qquad (12)$$

Заметим, что на последнем акте при кодировании (j,k)-й точки все ее окрестные точки принадлежат Ω_8, и, следовательно, в числителе вероятности принимают значения 1 или 0 в зависимости от реализации на них. Если пренебречь величиной $\lambda\hbar$ в числителе (12), то вероятность совпадения значения в (j, k)-й точке хотя бы с одним из значений в ее окрестности будет равна

$$\alpha = 4/(4+\lambda) \approx 1 - \lambda/4. \qquad (13)$$

Таким образом, λ принимает простой физический смысл: $\lambda/4$ определяет вероятность несовпадения значений в (j, k)-й точке ни с одним из значений в ее окрестности. В примере, рассмотренном ранее, $\lambda = 1/625$, и вероятность несовпадения должна быть оценена величиной 1/2500. Однако вследствие влияния ложных градаций вероятность несовпадения была равна 0,108 (см. [6]). Таким образом, игнорируя несовпадение, второй кодер практически во всех случаях проводит исправление ложных градаций и только с вероятностью $\lambda/4$ дает пропуск градации. Информационные потери могут быть представлены в виде

$$\Delta I = -(\lambda/4)\log(\lambda/1024) = (\lambda/4)(\log 1024 - \log\lambda) = \\ = (5/2)\lambda - (1/4)\lambda\log\lambda. \qquad (14)$$

В рассмотренном нами примере информационные потери не превышают 0,02 бит/элемент, что много меньше искажений, вносимых ложными градациями. Таким образом, введение второго кодера на последних актах передачи следует признать целесообразным по двум причинам: это приводит к коррекции изображения и значительному сокращению длины передаваемого кода.

Определим контрастность изображения как минимальную величину разрыва при переходе от области с одним описанием к области с другим описанием. Для широкого класса изображений эти области можно представлять гладкими функциями с разрывами первого рода на границах. В этом случае при анализе значений на опорной четверке, если разность между максимальным и минимальным значениями не превышает приведенной контрастности (т. е. контрастности, деленной на расстояния между точками), следует переходить на нуль-битовую кодировку и далее сравнивать кодируемое значение со средним значением опорной четверки. Здесь предполагается гладкий участок, и мы пытаемся интерполировать функцию, сравнивая интерполированное значение с действительным. Такой способ дает увеличение частоты нуль-битовой кодировки и уменьшение частоты несовпадений, что приводит к увеличению плотности кодирования.

Из вышеописанного видно, что требования повышения коэффициента сжатия и одновременного улучшения качества изображений не являются противоречивыми при данном способе кодирования, так как удлинение кода происходит из-за размывания границ и влияния шумов.

Проведенные исследования показали, что плотности кодирования в значительной мере зависят от сложности и состояния изображения и находятся в пределах от 2 до 0,5 бит/элемент. На рисунке показано восстановленное изображение, представленное кодом 0,5 бит/элемент. Исходное изображение не приведено, так как оно практически не отличается от восстановленного.

ЛИТЕРАТУРА

1. Цифровое телевидение/Под ред. М. И. Кривошеева.— М.: Радио и связь, 1979.
2. Кодирование изображений.— ТИИЭР, 1980, т. 68, № 3.
3. Landau H. J., Slepian D. Some Computer Experiments in Picture Processing for Bandwidth Reduction.— The Bell System Techn. J., 1971, vol. 50, N 5.
4. Бокштейн И. М., Брауде-Золотарев Ю. М. Цифровая интерполяция при двумерном анализе и синтезе изображений.— Труды НИИР, 1980, № 4.
5. Старков М. А., Пинчук А. И. Эффективное кодирование изображений.— Техника средств связи. Техника телевидения, 1981, № 3.
6. Карпова О. М., Старков М. А. Информационные свойства изображений.— Автометрия, 1982, № 2.

Поступила в редакцию 30 октября 1981 г.

УДК 621.391

А. Н. САФРОНОВ, И. Н. ТРОИЦКИЙ, О. И. ХАРИТОНОВА

(Москва)

СИНТЕЗ АДАПТИВНЫХ АЛГОРИТМОВ ПАРАЛЛЕЛЬНОЙ КОРРЕКЦИИ ИСКАЖЕННЫХ ВОЛНОВЫХ ФРОНТОВ

По общему признанию зарубежных и отечественных исследователей применение когерентных активных оптических систем является в настоящее время наиболее перспективным направлением в решении задач улучшения качества изображений, формируемых в присутствии оптически неоднородных непоглощающих сред [1—4]. Преимущество таких систем заключается в их способности преддетекторной компенсации фазовых искажений, вносимых в полезный сигнал средой распространения. Для достижения этой цели в приемный тракт, содержащий формирующую оптику и регистрирующее устройство в плоскости изображения, вводят дополнительно активный оптический элемент (АОЭ), способный изменять пространственный профиль фазы приходящей волны. В качестве АОЭ могут использоваться брэгговские ячейки, сегментированные, монолитные и пленочные зеркала. Стратегия коррекции при этом всегда опирается на информацию, регистрируемую в плоскости образа. Обычно АОЭ состоит из совокупности примыкающих друг к другу отражающих поверхностей (сегментов), способных перемещаться друг относительно друга и тем самым образовывать в совокупности требуемый профиль всего зеркала.

Во всех системах подобного рода создание такого профиля основано на методе пробных возмущений. Это подразумевает внесение посредством АОЭ пробных фазовых приращений в отраженный от АОЭ сигнал в пределах некоторого сегмента с целью максимизации определенного интегрального критерия — функции резкости (ФР). Смещение сегмента, дающее максимум ФР, фиксируется в АОЭ с последующим переходом к другому сегменту зеркала с той же целью. Вся процедура осуществляется в течение временного интервала «замороженности» неоднородной среды τ_0. Такой простейший алгоритмический подход тем не менее связан с наибольшими техническими трудностями [5]. Уже при разумном числе сегментов ($N \sim 10^2$), оправдывающем применение активной оптики, требование на быстродействие АОЭ становится практически невыполнимым, поскольку время переключения одного сегмента при этом должно удовлетворять неравенству

$$\tau_n \ll \tau_0/Npl,$$

где p, l — количество фазовых дискретов и циклов подстройки соответственно. В частности, при $\tau_0 \sim 10^{-3}$ с, $N \sim 10^2$, $pl = 10$ имеем $\tau_n \ll 10^{-6}$ с. Кроме того, в рамках описанной методики остается невыясненной воз-