

О. В. ЦВЕТКОВ, С. В. КУЛЕШОВ

Санкт-Петербургский институт информатики и автоматизации РАН

**О КРИТЕРИЯХ МЕТРОЛОГИЧЕСКОЙ ОЦЕНКИ
ИНФОРМАЦИОННЫХ СООБЩЕНИЙ**

Сформулировано основное свойство передачи компрессированных аудиовизуальных данных. Введены метрологическая характеристика оценки качества представления компрессированных данных и понятие алгоритмической сложности шума.

Традиционно для метрологической оценки качества представления данных — информационных сообщений — используются энергетические критерии, заимствованные из аналоговой техники и общей теории связи. Основными критериями по-прежнему являются среднеквадратичное отклонение и пиковое отношение сигнала к шуму. Эти критерии не вполне оправданы с точки зрения психофизиологической модели восприятия аудиовизуальных сигналов (далее — просто „сигналов“) и по этой причине не подходят для метрологической оценки параметров цифровых каналов передачи сигналов, в которых используются устройства компрессии.

В цифровых каналах можно выделить два принципиально различающихся типа процедур компрессии данных. Первый тип — компрессия без потерь — основан на использовании методов кодирования источника, которые могут строиться на принципах статистического (код Хаффмана и др.) или алгоритмического (программируемые технологии) снижения избыточности исходного цифрового представления. Второй тип — компрессия с потерями, при которой по каналу передается приблизительная копия исходного сигнала. При этом предполагается, что внесенные изменения не существенны при передаче информационного содержания, вследствие чего данный тип можно назвать *семантической компрессией* источника [1]. Метрологические аспекты этого типа компрессии являются предметом настоящей статьи.

Рассмотрим процедуру семантической компрессии сигнала с формальной точки зрения. Логика построения любого метода компрессии заключается примерно в следующем. Определяются закономерности, присущие внутренним связям между структурными элементами сигнала (отсчетами сигнала или его спектра, последовательными кадрами и т.п.), на основании которых можно предсказать значения одних элементов по известным значениям других. При этом, как правило, предсказание получается приближенным, кроме того, чем больше элементов сигнала берется в качестве исходных данных, тем точнее предсказание значений оставшихся. Элементы, взятые в качестве исходных, передаются по каналу связи, оставшиеся — предсказываются на основании найденных закономерностей. Таким образом, если на передающей стороне канала имеем исходный сигнал S , то на приемной — реконструированный сигнал \tilde{S} , определяемый как

$$\tilde{S} = S + N,$$

где N — шум, обусловленный компрессией сигнала.

Эта ситуация напоминает типовую модель аналогового канала, где шум канала складывается с переданным сигналом. Однако в цифровом канале имеется принципиальное отличие: все искажения сигнала происходят в процессе кодирования источника на передающей стороне, никаких искажений при передаче не происходит. Поэтому для цифрового канала более адекватной будет модель, рассматриваемая далее.

ближайшим? Наиболее распространенный ответ — по критерию минимума среднеквадратичного отклонения — не соответствует модели восприятия аудиовизуальной информации человеком, поэтому на данном этапе ограничимся ответом: по критерию минимума некоторой функции расстояния $d(S, \tilde{S})$. (Определение этой функции является предметом исследования следующей статьи, планируемой к публикации в одном из номеров журнала.)

При наличии исходного незашумленного изображения путем вычитания значений яркости соответствующих пикселей можно получить разностную картину, содержащую выделенный шум. Величина PSNR (peak signal-to-noise ratio) — пиковое отношение сигнала к шуму — обозначает соотношение между максимумом возможного значения сигнала и мощностью искажающего его шума. Эта величина наиболее часто используется для измерения уровня искажений при компрессии изображений. Принято определять PSNR через среднеквадратичное отклонение (СКО), которое для двух монохромных изображений I и J размером $m \times n$ (одно из них считается зашумленным приближением другого) вычисляется следующим образом (см.: <<http://en.wikipedia.org/wiki/PSNR>>):

$$\text{СКО} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} |I(i, j) - J(i, j)|^2;$$

PSNR определяется так:

$$\text{PSNR} = 10 \lg \left(\frac{Y_{\max_I}^2}{\text{СКО}} \right) = 20 \lg \left(\frac{Y_{\max_I}}{\sqrt{\text{СКО}}} \right),$$

где Y_{\max_I} — максимальное значение яркости пиксела; при разрядности пиксела 8 бит $Y_{\max_I} = 255$.

Для цветных изображений с тремя компонентами RGB на пиксел применяется такое же определение величины PSNR, но СКО вычисляется по всем трем компонентам (и делится на утроенный размер изображения).

Необходимость введения иных критериев для оценки качества представления информационных сообщений, компрессированных с использованием цифровых методов, можно показать на следующем примере.

Стандартное устройство для проверки видеотрактов — генератор цветовых полос (ГЦП) — работает по принципу синтеза аналогового сигнала требуемой формы. Все искажения, вносимые в выходной сигнал ГЦП, имеют аналоговую природу — отклонения формы или значений сигнала — и характеризуются наличием шума. Вместе с тем существует и принципиально иной способ формирования сигнала шаблона, заключающийся в том, что цифровая память содержит представление всех пикселей изображения, которое с помощью видеокодера переводится в аналоговый вид. В этом случае могут иметь место и другие искажения, например, отклонения цифрового шаблона от истинного, которые не могут быть исправлены аналоговым трактом. Подобные искажения наблюдаются при любой цифровой обработке изображений.

Для оценки информационных сообщений введем *критерий качества изображения, содержащего шум*. Под этим критерием, учитывающим психофизиологический фактор, понимается алгоритмическая сложность шума на разностном кадре (сложность по Колмогорову или энтропийная сложность). Энтропийную сложность шума экспериментально можно оценить через степень компрессии разностной картины при энтропийном кодировании (отношение размеров сжатого и несжатого файлов):

$$K = S_{\text{исх}} / S_{\text{к}}, \quad (5)$$

где $S_{\text{исх}}$ — битовый объем данных исходного изображения разностной картины, содержащей шум; $S_{\text{к}}$ — то же, после использования энтропийного сжатия.

Введенный критерий в ряде случаев оказывается более корректным по сравнению с критериями СКО и PSNR. Пусть искажение изображения характеризуется изменением нулевого уровня сигнала (равномерным увеличением значений яркости всех пикселей на одинаковую величину). В этом случае разностная картина будет представлять собой пустое изображение, имеющее высокую степень сжатия K согласно уравнению (5), что соответствует высокому качеству изображения. Это подтверждается экспериментальными данными — степень идентификации объекта на изображении практически не зависит от нулевого уровня сигнала, тогда как показатели СКО и PSNR будут непосредственно зависеть от этого вида искажений.

Полученные результаты позволяют изменить подход к организации цифровых каналов связи, а также ввести критерии для сравнения уже имеющихся каналов передачи данных.

СПИСОК ЛИТЕРАТУРЫ

1. Цветков О. В. Цифровые технологии обработки аудио-, видеосигналов: компрессия и семантический анализ // Тр. СПИИРАН. СПб.: СПИИРАН, 2004. Вып. 2.
2. Колмогоров А. Н., Успенский В. А. К определению алгоритма // Успехи мат. наук. 1958. Т. 13, вып. 4(82). С. 3—28.
3. Александров В. В., Горский Н. Д. Представление и обработка изображений. Рекурсивный подход. Л.: Наука, 1985. 190 с.
4. Alexandrov V. V., Gorsky N. D. From Humans to Computers: Cognition through Visual Perception. Singapore: World Scientific, 1991.