

Организация удаленного доступа к томографическим данным по медленному каналу связи¹

Д. В. Сушко, Ю. М. Штарьков

Институт проблем передачи информации, Российская академия наук, Москва, Россия

Поступила в редколлегию 15.11.2007

Аннотация—Рассмотрена задача организации удаленного доступа к томографическим данным, полученным в результате исследований на компьютерном томографе, по медленному каналу связи. Предложенная стратегия доступа вместе с разработанным и реализованным методом сжатия томографических данных позволяет обеспечить пользователям (врачам, исследователям и др.) комфортную работу с удаленными томографическими данными, например с использованием глобальной сети Интернет. Это, в свою очередь, открывает широкие возможности по использованию данных томографических исследований для проведения медицинских консультаций “на расстоянии”, видеоконференций, организации сетевых семинаров повышения квалификации медицинского персонала и т.д.

ВВЕДЕНИЕ

В настоящее время созданы и очень быстро пополняются многочисленные банки данных томографических исследований, широко используемые в медицине для диагностики, планирования лечения и т.д. Помимо долгосрочного хранения томографических данных (ТД), важно обеспечить возможность удаленного доступа к ТД. Последнее позволит пользователям (врачам, исследователям и др.) работать с удаленными ТД, проводить “на расстоянии” медицинские консультации, видеоконференции, семинары повышения квалификации медицинского персонала и т.д. Решение перечисленных задач затрудняется большим объемом ТД и выдвигает на первый план задачу их сжатия.

Основной задачей настоящей работы является задача организации удаленного доступа к ТД с использованием медленного канала связи. Принципиальное требование заключается в том, что ТД должны передаваться без искажений. В работе рассматриваются данные, полученные в результате проводимых на компьютерном (рентгеновском) томографе исследований. Такие данные представляют собой набор (от нескольких до нескольких сотен) полутоновых изображений (томограмм), каждое из которых описывает плотность тканей пациента в плоскости одного среза (слайса). Томограмма содержит 512×512 элементов, целочисленные яркости которых соответствуют относительным плотностям тканей в единицах Хаунсфилда и принимают значения в диапазоне $[-1024, 3071]$, т.е. описываются 12 битами. (Заметим, что обычно для хранения каждого пикселя томограммы используются два байта, т.е. 16 бит.) Таким образом, на передачу одной томограммы по сравнительно медленному каналу (со скоростью передачи около 1 МБт/с) затрачивается около 3.2 с, а на весь набор ТД — значительно больше. Такие задержки обмена данными неудобны для пользователя: известно, что работа в интерактивном режиме перестает быть комфортной, если время отклика на простые действия типа *Open* превышает 0.5 с. Один из основных способов уменьшения задержек — сжатие данных. Однако

¹ Работа частично поддержана в рамках программы фундаментальных исследований Президиума РАН “Фундаментальные науки — медицине”, проект “Разработка методов, алгоритмов и программ сжатия томографических изображений для их архивации, хранения и передачи по каналам связи”.

на практике, оставаясь в рамках требования сжатия данных без искажений, редко удается обеспечить степень сжатия изображений, большую 3 (см., например, [1]). В случае рентгеновских томограмм младшие 4 бита изображения содержат почти чистый белый шум [2], что связано с квантовой природой томографических измерений. Тем самым значение 3 может, с необходимыми оговорками, рассматриваться как верхняя оценка для коэффициента сжатия, а время доступа не может быть сделано меньше 1 с. Это означает, что для решения поставленной задачи необходимо, помимо сжатия, использовать разумную стратегию доступа, которая позволила бы сократить субъективное время отклика. Полное время доступа при этом, разумеется, не изменится.

Работа имеет следующую структуру. В разделе 1 рассматривается стратегия доступа, позволяющая сократить субъективное время отклика. В разделе 2 рассмотрены преобразования, уменьшающие диапазон значений исходных данных и согласующиеся с предложенной стратегией доступа. Метод кодирования рассмотрен в разделе 3. В заключении обозначены перспективные задачи по сжатию ТД.

1. СТРАТЕГИЯ ДОСТУПА

В основе предлагаемой стратегии лежат следующие естественные предположения относительно потребностей пользователя (медика) при работе с ТД.

- Пользователь не нуждается сразу во всех данных, в каждый конкретный момент времени ему необходима лишь та часть данных, которая содержит информацию об интересующей его части объекта (участке тела, мозга и т.п.). Такая информация содержится в небольшом наборе соседних томограмм, общее число которых редко превышает 5.
- Диапазон томограмм, охватывающих область интереса, как правило, известен а priori. В противном случае он может быть установлен при беглом просмотре ТД, для чего достаточно сформировать и предоставить пользователю “иконки” томограмм небольшого размера, которые для передачи по каналу связи могут кодироваться, например, JPG’ом. Все требуемые операции осуществляются в реальном времени.
- Детальное рассмотрение в любом случае начинается с некоторой (опорной) томограммы из области интереса и занимает время, заведомо превышающее 3 с.
- Обычно после детального рассмотрения указанной томограммы пользователю требуется рассмотреть также ближайшие к ней томограммы, отвечающие области интереса.

Предположим, что нам удалось разделить составляющие изображение данные на две части (объемы которых находятся в отношении $\sim 1:3$), так что меньшая часть (передаваемая в первую очередь) позволяет восстановить хорошее приближение исходного изображения, а оставшаяся часть (передаваемая вслед за первой) позволяет восстановить его полностью (без искажений). Тогда появляется возможность использовать приближение вместо оригинала в тот период времени, когда первая часть данных уже получена, а вторая — еще нет, и тем самым уменьшить время отклика в 4 раза. Если, кроме того, обеспечить сжатие каждой части данных не менее чем в два раза, то время отклика не будет превышать 0.4 с, что вполне приемлемо. Заметим, что как только вторая часть данных получена (приблизительно через 1.2 с после получения первой части), приближение заменяется оригиналом. После этого целесообразно немедленно начать передачу данных, отвечающих соседним томограммам из области интереса, так что если пользователь после просмотра первой томограммы затребует эти данные (что весьма вероятно), то они уже будут находиться на его локальном компьютере и время ожидания будет пренебрежимо мало.

Описанная выше стратегия позволяет обеспечить время отклика порядка 0.4 с для доступа к первой томограмме из области интереса и почти мгновенный доступ ко всем последующим

томограммам из той же области и, таким образом, решает поставленную задачу. Ключевыми моментами в реализации предложенной стратегии являются: 1) построение преобразования для разделения данных на части, 2) разработка алгоритма сжатия полученных в результате такого преобразования данных. Алгоритм сжатия должен обеспечить степень сжатия порядка 2. Время, необходимое на кодирование/декодирование данных, должно быть мало по сравнению со временем передачи данных по каналу связи.

2. ПРЕОБРАЗОВАНИЕ ИСХОДНЫХ ДАННЫХ

Любой метод сжатия изображений в той или иной форме использует статистическую зависимость между элементами изображения. Один из способов заключается в том, что исходные данные подвергаются некоторому преобразованию, которое, используя статистическую зависимость (корреляцию) элементов изображения, сужает диапазон возможных значений и/или изменяет распределение частот этих значений. (Иногда говорят, что такое преобразование уменьшает межэлементную избыточность на изображении [1].) К числу наиболее эффективных и успешно применяемых преобразований относятся дискретное косинусное преобразование, дискретные вейвлет-преобразования (по разным системам вейвлетов) и предсказания (обычно низкого порядка). В нашем случае искомое преобразование должно дополнительно обладать следующими свойствами: 1) быть обратимым, чтобы обеспечить сжатие без искажений, 2) не иметь большой вычислительной сложности и, наконец, 3) обеспечивать разделение данных на две части так, чтобы первая часть позволяла восстановить хорошее приближение оригинала. (Отметим, что косинусное преобразование не является точно обратимым из-за возникающих ошибок округления и поэтому неприменимо.) Всем указанным условиям полностью удовлетворяет дискретное вейвлет-преобразование (ДВП) на основе системы вейвлетов точного восстановления. В результате применения ДВП изображение оказывается разбитым на четыре равные части, первая из которых представляет собой приближение исходного изображения вдвое меньшего разрешения, а остальные три содержат “поправки”, позволяющие восстановить исходное изображение. Диапазон значений этих “поправок” значительно уже, чем у исходного изображения, а распределение частот их значений неравномерно.

Напомним, что представляет собой ДВП изображения (подробно о вейвлетах и их применении в обработке сигналов и изображений см. в книгах [3], [4]). В одномерном случае прямое ДВП может рассматриваться как фильтрация (свертка) исходной последовательности $\{x_n\}$ с низкочастотным $\{h_n^0\}$ и высокочастотным $\{h_n^1\}$ фильтрами анализа и последующим прореживанием, которое заключается в отбрасывании нечетных членов последовательностей, полученных после фильтрации:

$$x_n^i = \sum_k h_{2n-k}^i x_k, \quad i = \{0, 1\}. \quad (1)$$

(Индексы принимают все целые значения, и удобно считать, что конечные последовательности дополнены нулями). Таким образом, на выходе прямого преобразования мы получаем две последовательности $\{x_n^0\}$ и $\{x_n^1\}$, которые называются приближением и детальной составляющей исходной последовательности. Вид входящих в формулу (1) фильтров, которые задают конкретное преобразование, определяется выбором системы вейвлетов (по поводу связи ДВП с разложением функции по базису вейвлетов и кратномасштабным анализом см. [3]). Обратное ДВП представляет собой разбавление обеих последовательностей $\{x_n^0\}$ и $\{x_n^1\}$ нулями (т.е. построение новых последовательностей с нулевыми нечетными членами и четными членами, заданными последовательностями $\{x_n^0\}$ и $\{x_n^1\}$), фильтрацией (сверткой) полученных последовательностей с фильтрами синтеза $\{g_n^0\}$, $\{g_n^1\}$ и сложением результатов:

$$\hat{x}_n = \sum_k g_{n-2k}^0 x_k^0 + \sum_k g_{n-2k}^1 x_k^1. \quad (2)$$

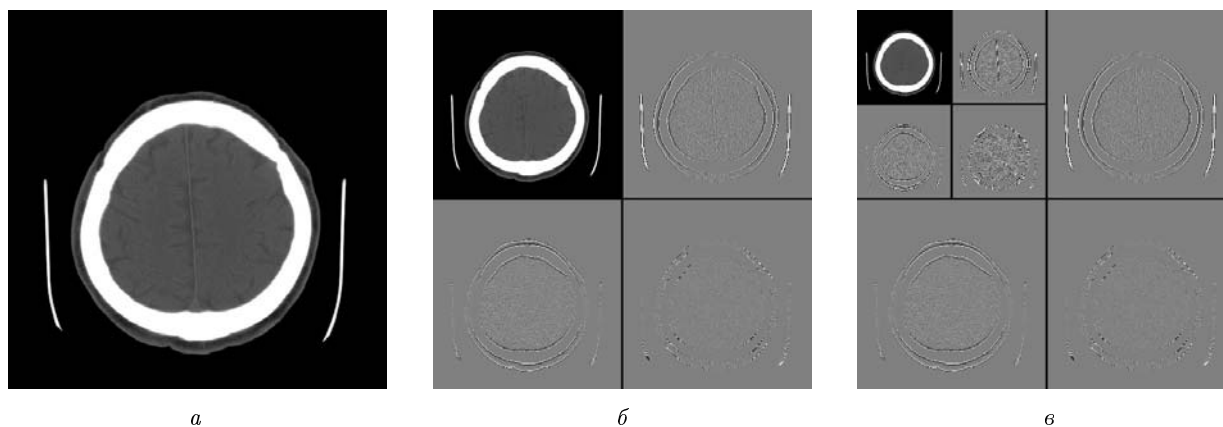


Рис. 1. Двумерное дискретное вейвлет-преобразование (ДВП): а) исходное изображение; б) двухмасштабное ДВП; в) трехмасштабное ДВП.

Чтобы преобразование (2) действительно было обратным по отношению к преобразованию (1) (т.е. $\{x_n\} \equiv \{\hat{x}_n\}$), фильтры анализа и синтеза должны удовлетворять следующим условиям:

$$\begin{cases} G^0(z)H^0(z) + G^1(z)H^1(z) = 2, \\ G^0(z)H^0(-z) + G^1(z)H^1(-z) = 0, \end{cases} \quad (3)$$

где H^0 , H^1 , G^0 и G^1 — символы последовательностей $\{h_n^0\}$, $\{h_n^1\}$, $\{g_n^0\}$ и $\{g_n^1\}$ (символом последовательности $\{a_n\}$ называется формальный ряд вида $A(z) \doteq \sum_n a_n z^{-n}$). Количество операций, требуемых для вычисления ДВП последовательности конечной длины N при использовании коротких фильтров (т.е. фильтров, у которых лишь небольшое число членов отлично от нуля), пропорционально N , причем значение соответствующей константы невелико. Тем самым процесс вычисления ДВП не требует больших временных затрат. Кроме того, очевидно, что последовательности, получаемые в результате прямого преобразования, имеют длину порядка $N/2$. Небольшая модификация изложенной схемы для аккуратного учета краевых эффектов приводит к тому, что приближение $\{x_n^0\}$ имеет длину, в точности равную $[N/2] + 1$, а детальная составляющая $\{x_n^1\}$ — $[N/2]$. Здесь и далее $[\cdot]$ — целая часть числа.

В случае изображений (двумерный случай) преобразование представляет собой суперпозицию одномерных преобразований, применяемых отдельно к строкам и столбцам. Пусть мы имеем изображение I размером $M \times N$ (для простоты будем считать, что число строк M и число столбцов N четные). Применяя прямое одномерное ДВП к каждой строке исходного изображения, мы получаем два промежуточных изображения I^0 и I^1 , размерами $M \times N/2$ каждое. Применяя теперь прямое одномерное ДВП к каждому столбцу изображений I^0 и I^1 , получаем окончательно четыре изображения I^{00} , I^{01} , I^{10} и I^{11} , размерами $M/2 \times N/2$, первое из которых называется приближением масштаба 1, а остальные — детальными составляющими масштаба 1 (соответственно горизонтальными, вертикальными и диагональными составляющими). Полученное разложение в целом называется двухмасштабным ДВП исходного изображения. Обратное преобразование совершается с использованием обратных одномерных ДВП, применяемых в обратном порядке (сначала к столбцам, затем к строкам).

На Рис. 1а представлена томограмма мозга, а на Рис. 1б — результат применения ДВП к данному изображению (двухмасштабное ДВП). (В данном примере при вычислении ДВП были использованы фильтры, соответствующие системе вейвлетов Легалла–Табатабаи, см. далее.) Левая верхняя четверть изображения на Рис. 1б представляет собой приближение масштаба 1, а оставшиеся части суть детали (детальные составляющие) масштаба 1. Уменьшение

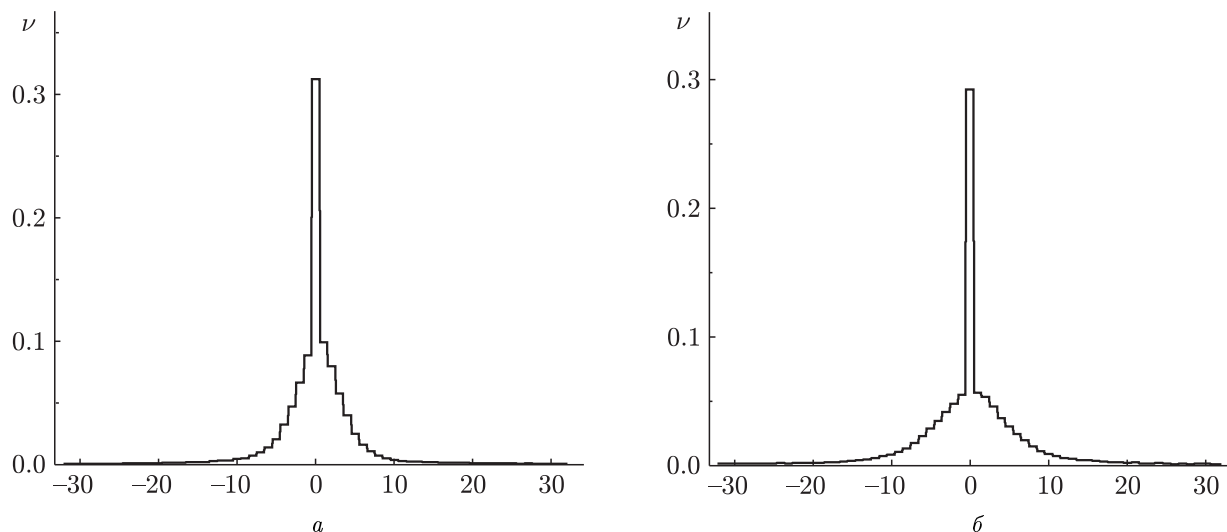


Рис. 2. Гистограмма распределения значений а) детальных составляющих масштаба 1, б) разности с предсказанием для приближения масштаба 1.

межэлементной избыточности для детальных составляющих очевидно: эффективный динамический диапазон значений детальных составляющих значительно меньше, чем на исходном изображении (на рисунке в иллюстративных целях отвечающие детальному составляющим части изображения были подвергнуты градиационному преобразованию, без этого различия в значениях яркости деталей была бы практически незаметна). Гистограмма распределения значений детальных составляющих масштаба 1 приведена на Рис. 2 а (по оси ординат отложены нормированные частоты). Для рассматриваемого примера оценка энтропии нулевого порядка (квазиэнтропия) для исходного изображения равна 7.42, а аналогичные оценки для приближения и детальных составляющих масштаба 1 равны соответственно 7.63, 4.58, 4.20 и 3.45 бит/пиксель. Напомним, что оценка энтропии нулевого порядка вычисляется по формуле

$$E_0 = - \sum_k \nu_k \log_2 \nu_k,$$

где суммирование ведется по всем встречающимся на изображении значениям, ν_k — нормированная частота k -го значения, и характеризует минимальную скорость кода, которая может быть достигнута при кодировании источника как источника без памяти.

В принципе возможно дальнейшее итеративное продолжение процедуры. Именно, повторное применение ДВП к приближению масштаба 1 (левая верхняя четверть изображения на Рис. 1 б) дает разложение, представленное на Рис. 1 в (трехмасштабное ДВП). В левой верхней четверти этого изображения располагаются приближение масштаба 2 и детальные составляющие масштаба 2. Все сказанное при рассмотрении двухмасштабного разложения может быть повторено дословно. Однако, как показали проведенные исследования, целесообразно ограничиться двухмасштабным разложением. Использование разложения с бóльшим числом масштабов было бы оправданно лишь в том случае, если бы это способствовало заметно более сильной степени сжатия. Этого, однако, не наблюдается.

Как было отмечено выше, конкретный вид блоков фильтров анализа и синтеза зависит от выбора системы вейвлетов, которая, в свою очередь, должна выбираться исходя из характера решаемой задачи. В том случае, когда необходимо обеспечить сжатие без искажений, нужно использовать системы вейвлетов, обеспечивающие точную обратимость преобразований (отсут-

ствие ошибок округления, см. далее). Такими системами являются наиболее простая система ортогональных вейвлетов Хаара, система биортогональных вейвлетов Легалла–Табатабаи [5], а также ряд других более сложных систем (см. [6, 7, 8]).

Для решения рассматриваемой задачи нами была выбрана система вейвлетов Легалла–Табатабаи. Ненулевые коэффициенты фильтров для этой системы принимают значения

$$\begin{aligned} h_{-2}^0 &= -1/8, & h_{-1}^0 &= +1/4, & h_0^0 &= +3/4, & h_1^0 &= +1/4, & h_2^0 &= -1/8, \\ h_{-2}^1 &= -1/2, & h_{-1}^1 &= +1, & h_0^1 &= -1/2, & & & & \\ g_{-1}^0 &= +1/2, & g_0^0 &= +1, & g_1^0 &= +1/2, & & & & \\ g_{-1}^1 &= -1/8, & g_0^1 &= -1/4, & g_1^1 &= +3/4, & g_2^1 &= -1/4, & g_3^1 &= -1/8. \end{aligned} \quad (4)$$

В том, что условия (3) выполнены, нетрудно убедиться непосредственно. Как показали проведенные эксперименты, данная система выгодно сочетает простоту и эффективность реализации и, в случае томографических данных, обеспечивает значительное сужение диапазона значений детальных составляющих, что демонстрирует, например, Рис. 2 а.

Обсудим проблему реализации ДВП. Преобразование должно быть точно обратимым, поэтому при его реализации должна быть обеспечена необходимая точность вычислений. Если прямое преобразование вычисляется непосредственно по формуле (1) с фильтрами (4), то для обеспечения точности вычислений мы вынуждены добавить три дополнительных бита для хранения значений приближения и один бит для детальной составляющей. Значения, попадающие в эти дополнительные “младшие” биты, представляют собой практически белый шум, что крайне нежелательно с точки зрения дальнейшего кодирования. Решение проблемы заключается в том, чтобы реализовать ДВП в рамках предложенной Малла так называемой лифтинг-схемы, что позволяет использовать целую арифметику. Кратко рассмотрим здесь эту схему применительно к выбранной системе вейвлетов (4) (подробнее см., например, [9]).

Рассмотрим снова одномерную бесконечную последовательность $\{x_n\}$ и применим к ней последовательно два преобразования:

$$x \rightarrow \tilde{y} : \quad \tilde{y}_{2n} = x_{2n}, \quad \tilde{y}_{2n+1} = x_{2n+1} - \frac{x_{2n} + x_{2n+2}}{2}; \quad (5)$$

$$\tilde{y} \rightarrow y : \quad y_{2n} = \tilde{y}_{2n} + \frac{\tilde{y}_{2n-1} + \tilde{y}_{2n+1} + 2}{4}, \quad y_{2n+1} = \tilde{y}_{2n+1}. \quad (6)$$

Прямое вычисление показывает, что $y_{2n} = x_n^0$ и $y_{2n+1} = x_n^1$, где $x_n^{0,1}$ вычисляются по формулам (1) с фильтрами (4). Таким образом, прямое ДВП может быть получено как суперпозиция преобразований (5) и (6). Аналогично проверяется, что обратное ДВП может быть получено как суперпозиция следующих преобразований:

$$y \rightarrow \tilde{y} : \quad \tilde{y}_{2n} = y_{2n} - \frac{y_{2n-1} + y_{2n+1} + 2}{4}, \quad \tilde{y}_{2n+1} = y_{2n+1}; \quad (7)$$

$$\tilde{y} \rightarrow x : \quad x_{2n} = \tilde{y}_{2n}, \quad x_{2n+1} = \tilde{y}_{2n+1} + \frac{\tilde{y}_{2n} + \tilde{y}_{2n+2}}{2}. \quad (8)$$

Теперь остается заметить, что если в формулах (5)–(8) округлять результаты деления до ближайшего целого снизу, то модифицированные таким образом прямое ДВП (5), (6) и обратное ДВП (7), (8) будут в точности обратными друг по отношению к другу. При этом каждое из преобразований отличается от своего немодифицированного аналога на величину, не превышающую ошибки округления (одной градации). При такой реализации требуется лишь один дополнительный бит для хранения знака.

Итак, после ДВП исходные данные оказываются естественным образом разбитыми на две части: приближение и детальные составляющие. Данные, соответствующие детальным составляющим, могут быть поданы на вход кодера непосредственно. В то же время статистические

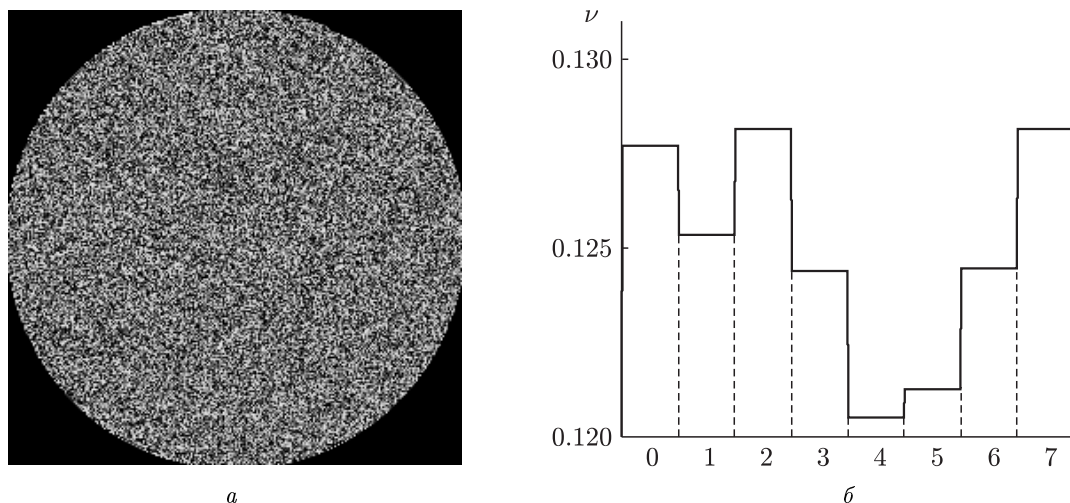


Рис. 3. а) Изображение 3 младших бит разности с предсказанием для приближения масштаба 1; б) гистограмма распределения значений нетривиальной части этого изображения.

характеристики приближения в целом мало отличаются от статистических характеристик исходного изображения. Для сужения эффективного диапазона значений этих данных можно использовать преобразование, состоящее в переходе от самих значений к значениям разности с некоторым предсказанием.

Для построения предсказания был разработан и реализован ряд алгоритмов нулевого, первого и второго порядков. Проведенные эксперименты показали, что в случае томограмм для приближения ДВП масштаба 1 наилучшие результаты достигаются при использовании простейшего предсказания нулевого порядка вида $\tilde{x}_{ij} = [(x_{i,j-1} + x_{i-1,j} + 1)/2]$, где x_{ij} и \tilde{x}_{ij} — значения изображения и предсказания в точке (i, j) , соответственно. (Предсказание текущего элемента равно полусумме значений верхнего и левого по отношению к нему элементов, округленной до ближайшего целого.) Гистограмма распределения значений разности с предсказанием для приближения масштаба 1 томограммы мозга (левая верхняя четверть на Рис. 1 б) приведена на Рис. 2 б. Соответствующая оценка энтропии нулевого порядка равна 6.00 бит/пиксель (напомним, что аналогичная оценка для самого приближения равна 7.63 бит/пиксель).

3. КОДИРОВАНИЕ ДАННЫХ

В рассматриваемом случае данные, которые предстоит кодировать, получены в результате преобразований, рассмотренных в предыдущем разделе. Первая часть данных (разность с предсказанием для приближения) описывается 15 битами, вторая часть (совокупность деталей составляющих) — 14 битами. Распределения частот значений обеих частей сконцентрированы вблизи нуля (см. Рис. 2). Анализ данных показывает, что 3–4 младших бита содержат практически белый шум. Сказанное иллюстрируется Рис. 3: на Рис. 3 а приведено изображение 3 младших бит разности с предсказанием для приближения масштаба 1 уже фигурирующей в примерах ранее томограммы мозга, а на Рис. 3 б — гистограмма соответствующих значений, вычисленная по нетривиальной части изображения.

Указанные особенности диктуют общую стратегию сжатия данных. Неравномерное кодирование можно осуществить по схеме, напоминающей аналогичную схему, принятую при JPEG кодировании [10]. Весь диапазон значений следует разделить на некоторое число неравных диапазонов. Диапазоны затем кодируются кодом неравномерной длины, а информация, опре-

деляющая конкретный символ внутри категории, описывается равномерным кодом. При этом соответствующие шуму младшие биты заведомо описывается равномерным кодом, так как любой другой способ кодирования шума может привести лишь к увеличению объема данных.

При таком подходе основной требующей решения задачей является конкретный выбор диапазонов и длин кодов для них. Важно решить вопрос о том, может ли такой выбор быть сделан раз и навсегда (как это, по существу, имеет место в большинстве случаев при JPEG кодировании) или необходимо использовать некоторый адаптивный (к входным данным) алгоритм выбора. При решении этого вопроса необходимо принять во внимание временные затраты, которые потребуются на реализацию того или иного адаптивного алгоритма. В отличие от процедуры декодирования, которая не занимает заметного (в сравнении со временем передачи по каналу) времени, достаточно сложный адаптивный алгоритм может потребовать значительного времени на реализацию. Поскольку это время входит составной частью в общее время доступа, оно заведомо не должно превышать ~ 0.2 с. С учетом сказанного на первом этапе разработки было принято решение отказаться от использования адаптивных алгоритмов и использовать постоянную кодовую таблицу. Таблица, приведенная ниже, используется в окончательном варианте разработанного кодека (значения длин приведены в битах).

Таблица

№	Диапазон значений	Длина кода диапазона	Длина описания	Длина кода
0	$[-4, -1] \cup [0, 3]$	1	3	4
1	$[-8, -5] \cup [4, 7]$	2	3	5
2	$[-16, -9] \cup [8, 15]$	3	4	7
3	$[-32, -17] \cup [16, 31]$	4	5	9
4	$[-64, -33] \cup [32, 63]$	5	6	11
5	$[-128, -65] \cup [64, 127]$	6	7	13
6	$[-256, -129] \cup [128, 255]$	7	8	15
7	$[-512, -257] \cup [256, 511]$	8	9	17
8	$[-1024, -513] \cup [512, 1023]$	9	10	19
9	остальные значения	9	14	23

Построенный кодек обеспечивает в среднем двукратное сжатие томографических данных. Приведем результаты для рассмотренной выше томограммы мозга. Размер исходных данных равен ~ 530 КВ. Размер данных, необходимых для восстановления приближения вдвое меньшего разрешения, равен ~ 54 КВ, а размер дополнительных данных, необходимых для точного восстановления, составляет ~ 116 КВ. Общий коэффициент сжатия, таким образом, в данном случае равен ~ 3.1 , а эффективный (с учетом того, что каждые два байта в исходном файле содержат лишь 12 бит информации) — ~ 2.3 .

ЗАКЛЮЧЕНИЕ

Предложенный в настоящей работе подход позволяет успешно решить основную поставленную задачу: обеспечить пользователю комфортную работу с компьютерными томограммами, находящимися в удаленном банке данных, даже при отсутствии быстрого канала связи. Описанный в разделах 2 и 3 алгоритм сжатия реализован в виде динамически присоединяемой библиотеки (DLL) для операционной системы WINDOWS. Вместе с тем, проведенные исследования позволяют сформулировать ряд задач, требующих решения.

Первоочередной задачей является переход к использованию более совершенных методов кодирования в алгоритме сжатия. Оценки показывают, что только за счет аккуратной реализации арифметического кодирования данных как источника без памяти можно увеличить

степень сжатия на $\sim 10\text{--}15\%$. Еще больший выигрыш может дать использование методов универсального кодирования сообщений с неизвестной статистикой.

Перспективное направление исследований связано с переходом от сжатия индивидуальной томограммы к сжатию всего трехмерного массива ТД. Набор томограмм, получаемый в рамках одного томографического исследования, характеризуется значительной “межкадровой” избыточностью: данные, отвечающие томограммам соседних срезов, коррелированы, причем тем больше, чем меньше эффективное расстояние между срезами. В настоящее время уже встречаются данные, для которых расстояние между соседними срезами и расстояние между соседними точками в пределах одной томограммы сравнимы по величине. В таком случае эффективное использование указанных статистических связей может привести к заметному увеличению степени сжатия всего массива ТД. Один из возможных подходов здесь может состоять в том, чтобы, рассматривая массив ТД как трехмерный массив, применить к нему трехмерное ДВП. Такой подход выглядит очень привлекательно и, по-видимому, никогда ранее не применялся.

При сжатии ТД без искажений нельзя рассчитывать на достижение в среднем коэффициента сжатия, существенно превышающего 2–3. Для достижения заметно лучшей степени сжатия нужно отказаться от требования полного сохранения информации, заменив его требованием сохранения существенной, т.е. медицинской, информации. В случае компьютерной томографии последнее требование допускает формализацию. Как известно, качество компьютерного томографа описывается набором метрологических характеристик, таких как уровень шума, функция передачи скачка, функция восстановления круга, кривая потери контраста, кривая контраст–размер, кривая передачи модуляции. Указанные метрологические характеристики вычисляются на основе томограмм определенных образцов (фантомов) по установленной методике (см., например, [11]). Метрология может быть взята за основу при определении понятия сжатия без существенных потерь (без потерь медицинской информации). Идея состоит в том, чтобы назвать некоторый алгоритм алгоритмом сжатия без существенных потерь, если метрологические характеристики, вычисленные по томограммам, которые были предварительно подвергнуты сжатию, не хуже исходных метрологических характеристик. Такой подход, не исключающий, разумеется, возможности субъективных оценок алгоритма со стороны экспертов и медиков, позволит объективно оценить качество алгоритмов сжатия с искажениями. Здесь, однако, возникает трудность следующего характера. В основе принятой метрологии лежит линейная модель томографии, которая в случае медицинской рентгеновской томографии хорошо отвечает существу дела. В то же время алгоритмы сжатия с искажениями, как правило, нелинейны (нелинейность возникает на этапе квантования данных). Поэтому реализация предложенной идеи потребует “правильного” обобщения метрологии, позволяющего учесть возникающую нелинейность.

СПИСОК ЛИТЕРАТУРЫ

1. Р. Гонсалес, Р. Вудс. *Цифровая обработка изображений*. М.: Техносфера, 2005.
2. Д. В. Сушко, Ю. М. Штарьков. О сжатии томографических данных. *Информационные процессы* (в печати).
3. I. Daubechies. *Ten Lectures on Wavelets*. Philadelphia: Society for Industrial and Applied Mathematics, 1992.
4. S. Mallat. *A Wavelet Tour of Signal Processing*. New York: Academic Press, 1999.
5. D. Le Gall and A. Tabatabai. Sub-band Coding of Digital Images Using Symmetric Short Kernel Filters and Arithmetic Coding Techniques. *IEEE International Conference on Acoustic, Speech, and Signal Processing*, New York, 1988, pp. 761–765.

6. A. R. Calderbank, I. Daubechies, W Sweldens, and B.-L. Yeo. Wavelet transforms that map integers to integers. *Applied and Computational Harmonic Analysis*, 1998, vol. 5, no. 3, pp. 332–369.
7. H. Chao, P. Fisher, and Z. Hua. An approach to integer wavelet transforms for lossless image compression. In: *Proc. of International Symposium on Computational Mathematics*, Guangzhou, China, 1997, pp. 19–38.
8. M. D. Adams and F. Kossentini. Reversible integer-to-integer wavelet transforms for image compression: Performance evaluation and analysis, *IEEE Trans. on Image Processing*, 2000, vol. 9, no. 6, pp. 1010–1024.
9. W. Sweldens. The lifting scheme: A custom-design construction of biorthogonal wavelets. *Applied and Computational Harmonic Analysis*, 1996, vol. 3, no. 2, pp. 186–200.
10. G. K. Wallace. The JPEG Still Picture Compression Standard. *Comm. ACM*, 1991, vol. 34, pp. 31–44.
11. D. A. Попов, Е. В. Sokolova, and D. V. Sushko. Mathematical Models in Two-Dimensional Radon Tomography. In: *Applied Problems of Radon Transform*, Ed. S. G. Gindikin, Amer. Math. Soc. Transl. (2), 1994, vol. 162, pp. 129–204.

Статью представил к публикации член редколлегии Н. А. Кузнецов