

# ИССЛЕДОВАНИЕ МЕТОДОВ СНИЖЕНИЯ ИНФОРМАЦИОННОЙ ИЗБЫТОЧНОСТИ ТЕЛЕВИЗИОННОГО СИГНАЛА И ОЦЕНКИ ИХ ЭФФЕКТИВНОСТИ

А. В. Скурихин, А. А. Бритиков

Санкт-Петербургский институт информатики и автоматизации РАН  
199178, Санкт-Петербург, 14-я линия В.О., д.39

<andrey@iias.spb.su>

---

УДК 681.3

А. В. Скурихин, А. А. Бритиков. **Исследование методов снижения информационной избыточности телевизионного сигнала и оценки их эффективности** // Труды СПИИРАН. Вып. 2, т. 1. — СПб.: СПИИРАН, 2004.

**Аннотация.** Приводится обзор различных подходов, направленных на снижение информационной избыточности телевизионного сигнала. Показываются преимущества предлагаемого метода адаптивно-динамической сегментации с точки зрения психофизиологического восприятия. Обсуждаются методы оценки эффективности снижения информационной избыточности. — Библ. 6 назв.

UDC 681.3

A. V. Skourikhin, A. A. Britikov. **The research of the methods for TV signal information redundancy decreasing and estimation of its efficiency** // SPIIRAS Proceedings. Issue 2, vol. 1.— SPb.: SPIIRAS, 2004.

**Abstract.** The different approaches to TV signal information redundancy decreasing are viewed. The advantages of proposing method of adaptive dynamic segmentation from the viewpoint of psychophysiological perception are shown. The methods for efficiency estimation of information redundancy decreasing are considering. — Bibl. 6 items.

---

## Введение

Телевизионный сигнал традиционно считается одним из наиболее сложных сигналов для задач анализа и передачи информации. Это обусловлено, в первую очередь, большим объемом содержащихся в нем данных, которые необходимо обработать и передать потребителю в реальном времени. С точки зрения классической теории информации, успешное решение этой задачи требует очень большой полосы пропускания по сравнению с другими сигналами, ориентированными на непосредственное восприятие человеком.

Широкое внедрение цифровых технологий найти решение проблемы снижения пропускной способности канала передачи телевизионного сигнала вне рамок традиционного подхода к передаче информации, ориентирующегося на непрерывные сигналы. Использование дискретного представления по Шеннону-Котельникову непосредственно само по себе не дает выигрыша в передаче информации, однако позволяет применить мощные методы цифровой обработки для снижения информационной избыточности и сокращения объема данных. Согласно исследованиям, сокращение избыточности телевизионного сигнала со статистической точки зрения не дает особых преимуществ с точки зрения пропускной способности.

Новая парадигма цифровой технологии передачи и обработки информации [1] подходит к проблеме информационной избыточности с принципиально иной точки зрения. В основе новой парадигмы лежат следующие основные положения:

— методы обработки, оперирующие с непрерывными величинами и бесконечными длинами, неадекватны с точки зрения цифровой технологии. Устройства цифровой обработки оперируют с дискретными наборами данных конечной длины, в общем случае — с двоичным цифровым потоком нулей и единиц;

— любой цифровой поток конечной длины можно заменить программой, порождающей этот поток с помощью элементарных машинных операций, манипулирующих с двоичными нулями и единицами. Прототипом системы, реализующей данные операции, может служить машина Тьюринга;

— представление данных в системе цифровой обработки телевизионных сигналов должно быть ориентировано на человека как конечного потребителя и на возможности его системы восприятия. Объекты, не воспринимаемые зрительной системой человека, являются избыточными. Человек обычно имеет дело с более крупными объектами, чем минимальный элемент цифрового изображения, который определяет разрешающую способность зрительной системы человека. Согласно психофизиологическим исследованиям, для человека характерно структурированное представление зрительных образов с рекурсивными обратными связями, необходимое для решения задач идентификации и узнавания.

В данной работе оценивается эффективность метода адаптивно-динамической сегментации, снижающего информационную избыточность телевизионного сигнала в рамках новой парадигмы цифровой обработки сигналов. Показываются преимущества данного метода по сравнению с методами, основанными на статистических критериях оценки избыточности и интегральных преобразований, берущих начало в области анализа и обработки непрерывных сигналов. Поскольку при предлагаемом подходе статистические методы типа СКО уже не могут служить адекватной мерой оценки, обсуждаются новые методы оценки эффективности снижения избыточности и перспективы практической реализации рассматриваемого метода.

## **1. Низкоуровневые (физиологические) особенности человеческого восприятия**

Особенности восприятия системы зрения человека традиционно подразделяются на две группы: низкоуровневые (физиологические) и высокоуровневые (психофизиологические) [2]. До недавнего времени большинство методов снижения информационной избыточности принимало во внимание низкоуровневые свойства системы человеческого зрения. В первую очередь это обусловлено тем, что свойства низкоуровневого восприятия хорошо описываются в рамках традиционных математических моделей, в частности, методами спектрального и вейвлет-анализа. Благодаря широкому распространению новых математических моделей, трактующих процесс восприятия и понимания аудиовизуального потока данных вне рамок интегрального представления, в последние годы наметилась тенденция к разработке методов и алгоритмов, учитывающих и высокоуровневые, психофизиологические характеристики системы зрительного восприятия человека.

Среди низкоуровневых особенностей человеческого восприятия видеоданных выделяют три наиболее важных свойства, влияющих на заметность постороннего шума в изображении: чувствительность к изменению яркости изображения, частотную чувствительность и эффект маскирования.

Чувствительность к изменению яркости можно определить следующим образом. Испытуемому показывают некоторую однотонную картинку (рис. 1). После того, как глаз адаптировался к ее освещенности  $I$ , «настроился на нее», постепенно изменяют яркость вокруг центрального пятна. Изменение освещенности  $\Delta I$  продолжают до тех пор, пока оно не будет обнаружено. Зависимость минимального контраста  $\Delta I/I$  от яркости  $I$ , показанная на рис. 1, получила название закона Вебера–Фехнера [3]. Как видно из рисунка, на большом диапазоне изменений яркости в области средних значений контраст примерно постоянен, тогда как для малых и больших яркостей значение порога неразличимости возрастает. Установлено, что  $\Delta I \approx 0.01 \div 0.03 \cdot I$  для средних значений яркости.

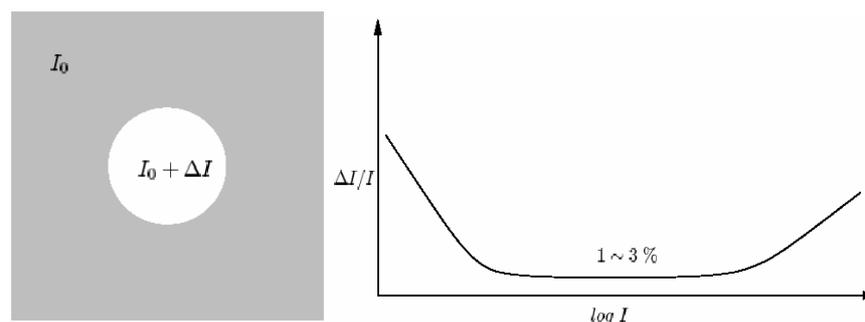


Рис. 1. Закон Вебера–Фехнера

Интересно заметить, что результаты новейших исследований противоречат «классической» точке зрения и показывают, что при малых значениях яркости для системы человеческого зрения порог неразличимости уменьшается, то есть зрение человека более чувствительно к шуму в этом диапазоне.

Частотная чувствительность зрительной системы проявляется в том, что человек гораздо более восприимчив к низкочастотному (НЧ), чем к высокочастотному (ВЧ) шуму. Это связано с неравномерностью амплитудно-частотной характеристики системы зрения человека. Экспериментально ее можно определить при помощи того же опыта, что и при яркостной чувствительности. Но на этот раз в центральном квадрате изменяются пространственные частоты до тех пор, пока изменения не станут заметными.

Элементы зрительной системы человека разделяют поступающий видеосигнал на отдельные компоненты. Каждая составляющая возбуждает нервные окончания глаза через ряд подканалов. Выделяемые глазом компоненты имеют различные пространственные и частотные характеристики, а также различную ориентацию (горизонтальную, вертикальную, диагональную) [3]. В случае одновременного воздействия на глаз двух компонентов со сходными характеристиками возбуждаются одни и те же подканалы. Это приводит к эффекту маскирования, заключающегося в увеличении порога обнаружения видеосигнала в присутствии другого сигнала, обладающего аналогичными характеристиками. Поэтому, аддитивный шум гораздо заметнее на гладких участках изображения, чем на высокочастотных, то есть в последнем случае наблюдается маскирование. Наиболее сильно эффект маскирования проявляется, когда оба сигнала имеют одинаковую ориентацию и местоположение.

Можно показать, что частотная чувствительность тесно связана с яркостной. Известно также и выражение для определения порога маскирования на основе известной яркостной чувствительности, что позволяет найти приемле-

мые (незаметные для глаза) искажения изображения, учитывающие свойства зрения человека. Такого типа математические модели хорошо разработаны для случая различных спектральных преобразований, в частности, для дискретного косинусного фурье-преобразования изображения. Между тем применение дискретного косинусного преобразования обусловлено в первую очередь широким распространением стандарта JPEG, который далеко не всегда эффективен по соотношению качество сжатия/субъективное восприятие, во многом уступая новому стандарту JPEG2000, основанному на вейвлет-преобразовании.

Эффект маскирования в пространственной области может быть объяснен путем построения традиционных (стохастических) моделей изображения. При этом изображение представляется в виде марковского случайного поля, распределение вероятностей которого подчиняется обобщенному гауссовскому закону.

Рассмотрим основные методы, учитывающие низкоуровневые особенности человеческого восприятия.

## 2. Фурье-анализ

Основные свойства фурье-анализа опишем для одномерных сигналов во временной области, однако все нижеизложенное распространяется и на многомерный фурье-анализ.

Как известно, интегральное преобразование Фурье и ряды Фурье являются основой гармонического анализа стационарных непрерывных сигналов. Получаемые в результате преобразования коэффициенты Фурье поддаются достаточно простой физической интерпретации, причем простота ни в коем случае не умаляет важности последующих выводов о характере исследуемого сигнала. Применение интегрального преобразования и рядов Фурье в вычислениях и аналитических преобразованиях очень наглядно, все необходимые свойства и формулы выписываются с помощью всего двух действительных функций  $\sin t$ ,  $\cos t$  (или одной комплексной — синусоидальной волны  $\exp(it) = \cos t + i \cdot \sin t$ ,  $i = \sqrt{-1}$ ) и достаточно легко доказываются.

Другой, не менее важной причиной популярности фурье-анализа явилось наличие адекватной аналоговой элементной базы (RLC-цепочек), позволившие в свое время совершить существенный прорыв в области передачи данных. Язык фурье-анализа является естественным языком описания функционирования данных устройств. Несмотря на существенное превосходство цифрового представления на данном уровне развития технических средств и явное несоответствие ключевых свойств целого ряда сигналов классической спектральной модели, ее популярность до сих пор остается преобладающей.

В классическом спектральном анализе сигнал во временной области удобно представлять в виде дискретизации по Шеннону, или разложения по  $\delta$ -функциям Дирака

$$f(t) = \int_{\mathbb{R}} f(u) \delta(t - u) du ,$$

и, соответственно, его частотное представление или преобразование Фурье — не что иное, как разложение по комплексным экспонентам

$$f(t) = \int_{\mathbb{R}} F(\omega) \exp(i\omega t) d\omega .$$

Первая формула дает информацию о максимально возможном разрешении по времени: величина  $f(t)$  представляет интенсивность сигнала в момент времени  $t$ . Никакая частотная информация при этом нам недоступна, простое поточечное представление не дает никакой информации о частотном содержании сигнала. Напротив, фурье-представление дает точное частотное представление, не обеспечивая никакой информации о временной локализации событий. На практике далеко не все сигналы стационарны. Пик в сигнале во временной области распространится во всей частотной области его преобразования Фурье. Можно провести аналогию с аудиозаписью. Ее временное представление позволяет нам знать, ноты или паузы мы слышим, но какую именно ноту, определить не позволяет. Напротив, частотное представление говорит, какие тоны присутствуют в музыке, не говоря ничего, в какой момент они могут быть услышаны.

Каждое из представлений содержит полную информацию о сигнале, поскольку преобразование Фурье является взаимно-однозначным отображением из временного представления в частотное и наоборот. Однако в каждом из случаев доступен только один вид интересующей нас информации. Иными словами, распределение по импульсам Дирака имеет сколь угодно малое разрешение по времени, в то время как в случае комплексных экспонент – по частоте. Вопрос состоит в нахождении представления сигнала, лежащего между этими крайностями, в котором дана смешанная информация об обоих этих параметрах: «в этот момент мы слышим ноту *до*, а здесь — ноту *фа*».

Первые попытки частичного устранения неспособности анализа Фурье осуществлять временную локализацию сингулярностей сигналов заключались в применении оконного преобразования Фурье. В стандартное преобразование вводится умножение на так называемое окно — движущуюся по временной оси функцию, имеющую компактный носитель. По сути дела, при оконном фурье-анализе мы разбиваем исходный сигнал на области одинакового размера и проводим спектральный анализ в каждой из этих областей. Именно на таком принципе основан стандарт сжатия изображений JPEG и близкие к нему стандарты сжатия видеоданных MPEG. Исходный кадр разбивается на непрерывающиеся блоки размером  $8 \times 8$  элементов. Далее каждый блок подвергается дискретному косинусному преобразованию Фурье. Несущественные для восприятия высокочастотные пространственные составляющие огрубляются и эффективно кодируются, обеспечивая для объектов с плавным изменением яркости экономичное представление.

Однако введение оконных функций имеет существенный недостаток, который состоит в том, что ширина окна при данном виде анализа фиксирована и поэтому не может быть адаптирована к локальным свойствам сигнала. Возникают и другие дополнительные проблемы, уже не поддающиеся решению в рамках традиционного фурье-анализа. К этим проблемам относится, в частности, мозаичная структура при больших коэффициентах сжатия, обусловленная независимой обработкой отдельных блоков изображения. В рамках фурье-анализа успешное решение проблемы оказалось невозможным из-за самой специфики интегральных преобразований, что заставило обратить внимание на методы, учитывающие локальные особенности сигналов и изображений.

### 3. Переход от фурье-анализа к вейвлет-анализу

Первый шаг к новому виду анализа сделал Габор, который предложил использовать базисные функции, лежащие посередине между импульсами Дирака и комплексными экспонентами, иными словами, локализованные как во времени, так и по частоте.

Согласно классическому принципу неопределенности, нельзя найти функцию, локализирующую событие в частотно-временной области с произвольно малой погрешностью. Однако известны функции, позволяющие провести локализацию с теоретически максимальной точностью. Это семейство смещенных во времени и модулированных гауссовских функций вида

$$g(t) = A \cdot \exp\left(-\frac{(t - t_0)^2}{2\Delta t^2}\right) \cdot \exp(i\omega_0 t),$$

где  $A$  — нормализующий коэффициент. Только для этих функций неравенство Гейзенберга обращается в равенство. Теперь эти функции называют вейвлетами Габора. Очевидно, что в случае  $t_0 = 0$ ,  $\omega_0 = 0$ ,  $\Delta t = 0$  габоровский вейвлет вырождается в функцию Гаусса, для которой, таким образом, также характерна минимальная частотно-временная неопределенность.

Сравним поведение габоровских вейвлетов в частотно-временной области с традиционными методами представления: дискретно-временным и спектральным. Частотно-временные покрытия для дискретно-временного представления по Шеннону, анализирующей функцией которого является функция Дирака, и для преобразования Фурье с комплексными экспонентами представляют собой прямоугольники бесконечно малой ширины и бесконечно большой длины, как показано на рис. 2.

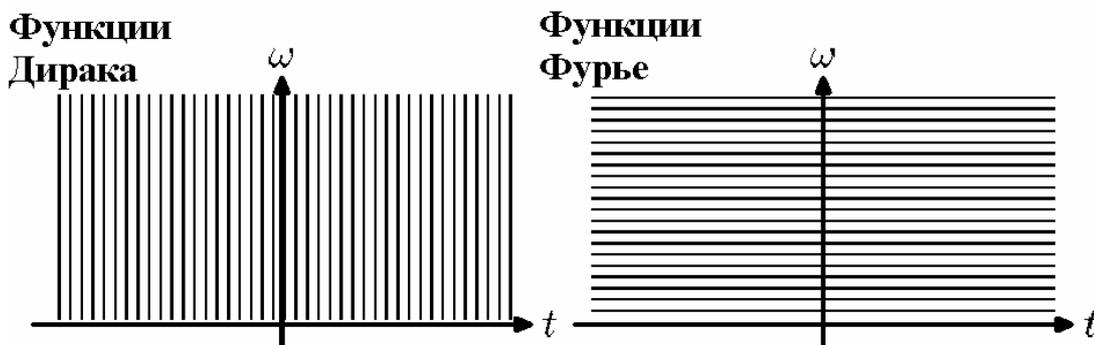


Рис. 2. Частотно-временное покрытие для представлений Фурье и Дирака

Очевидно, что представление любой функции в виде линейной комбинации обоих этих функций с использованием всех возможных комбинаций параметров  $t_0$ ,  $\omega_0$  и  $\Delta t$  будет явно избыточным. Для устранения этой избыточности были предложены два различных подхода:

— *Частотно-временной подход*, который состоит в выборе временной протяженности функций  $g$  независимо от частоты модуляции. Такое представление совпадает с упоминаемым выше оконным преобразованием Фурье.

— *Масштабно-временной подход*, в котором ширина функций во времени обратно пропорциональна их частоте, то есть произведение  $\omega_0 \Delta t$  есть величина постоянная.

Покрывания частотно-временной плоскости наглядно демонстрируют разницу между двумя этими подходами (рис. 3). В случае частотно-временного пред-

ставления покрытие состоит из одинаковых прямоугольников, транслируемых по всей плоскости. В случае же масштабно-временного представления прямоугольники имеют одинаковую площадь, но их относительная разрешающая способность по частоте остается при этом неизменной.

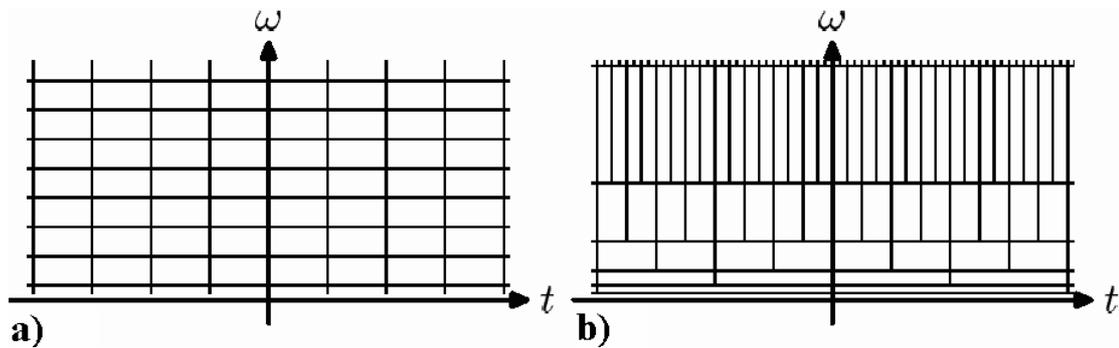


Рис. 3. Покрытие частотно-временной плоскости для частотно-временного (а) и масштабно-временного (б) представлений

Такое масштабно-временное представление положило начало новому способу анализа сигналов, именуемому вейвлет-анализом.

#### 4. Вейвлет-анализ

Вейвлет-анализ имеет дело с масштабно-временным представлением базисных функций. Этот подход имеет большое число преимуществ, главными из которых являются наличие быстрых алгоритмов вычислений и пропорциональность частотной и временной протяженности каждой функции к разрешающей способности по частоте и времени, определяемой принципом неопределенности [4].

Благодаря своей частотно-временной локализации вейвлет-функции способны аппроксимировать тонкую, нестационарную структуру сигналов. Локализация вейвлет-функций во времени подразумевает наличие концентрации их энергии внутри некоторого конечного интервала. Частотная локализация вейвлет-функций говорит о компактности носителя их фурье-образа, т.е. локализации его энергии внутри определенного частотного интервала.

Вейвлеты позволяют анализировать более тонкую структуру сигнала, подавляя медленно изменяющиеся его составляющие. При этом вейвлет-функции порождают ортонормированный базис, удовлетворяют концепции кратноразрешающего анализа, а также обеспечивают аппроксимацию данных с энтропией, меньшей исходной.

Масштабно-временное вейвлет-представление обычно использует растяжения и переносы во времени одной фиксированной вейвлет-функции  $\psi \in L^2(\mathbb{R})$ , имеющие вид

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right), \text{ где } a, b \in \mathbb{R}, a \neq 0.$$

При этом непрерывное вейвлет-преобразование функции  $f \in L^2(\mathbb{R})$  определяется по формуле

$$W_f(a, b) = \langle f(t), \psi_{a,b}(t) \rangle = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} f(t) \psi\left(\frac{t-b}{a}\right) dt.$$

По сути, прямое непрерывное вейвлет-преобразование представляет собой бесконечный набор линейных сверток анализируемой функции с вейвлет-функциями изменяющегося масштаба.

В случае дискретного вейвлет-преобразования применяется концепция кратноразрешающего (multiresolution) анализа, которая подразумевает представление пространства сигнала в виде иерархически вложенных подпространств и выборе в этих подпространствах ортонормированного вейвлет-базиса. Функции этого базиса имеют дискретный набор значений масштабирования, отличающихся на соседних уровнях иерархии в два раза, и дискретный набор значений сдвига, имеющих смысл дискретной выборки значений. При этом длина дискретного сигнала и разрешающая способность по времени, равная периоду дискретизации, делают возможным использование конечного набора базисных функций, соответственно ограничивая их масштабы сверху и снизу.

Метод устранения избыточности видеоданных на основе вейвлет-анализа описывается следующим образом. Для каждого кадра производится декомпозиция по пространственным составляющим, разделяя текущее изображение на сглаженную версию более грубого уровня иерархии и дифференциальные составляющие, необходимые для формирования более четкого уровня иерархии. В полученном пирамидальном представлении количество уровней иерархии определяется количеством дискретных элементов по ширине и высоте кадра. Пример вейвлет-декомпозиции показан на рис. 4.

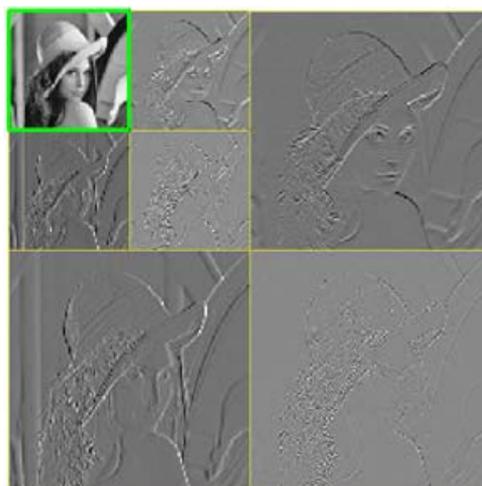


Рис. 4. Вейвлет-декомпозиция изображения

Дифференциальные составляющие данной пирамиды содержат малосущественные для низкоуровневого восприятия коэффициенты, которые подвергаются огрублению с последующим эффективным кодированием.

## 5. Высокоуровневые (психофизиологические) особенности человеческого восприятия

Высокоуровневые (психофизиологические) свойства системы человеческого зрения пока редко учитываются в методах уменьшения информационной избыточности. Их отличием от низкоуровневых является то, что эти свойства проявляются «вторично», то есть обработавший первичную информацию от зрительной системы мозг выдает команды на ее «подстройку» под изображение [3]. Основными из психофизиологических свойств являются:

— Чувствительность к контрасту. Высококонтрастные участки изображения, перепады яркости обращают на себя значительное внимание.

— Чувствительность к размеру. Большие участки изображения «заметнее» меньших размером. Причем существует порог насыщения, когда дальнейшее увеличение размера не существенно.

— Чувствительность к форме. Длинные и тонкие объекты вызывают большее внимание, чем круглые однородные.

— Чувствительность к цвету. Некоторые цвета (например, красный) «заметнее» других. Этот эффект усиливается, если фон заднего плана отличается от цвета фигур на нем.

— Чувствительность к местоположению. Человек склонен в первую очередь рассматривать центр изображения.

— Чувствительность к взаимному расположению объектов сцены. Зрительная система человека обычно акцентирует больше внимания на объектах переднего плана, чем заднего плана и фона.

— Чувствительность к контексту. Человеческое зрение в первую очередь акцентирует внимание на изображениях людей, а в пределах области лица выделяет глаза и рот.

— Чувствительность к внешним раздражителям. Движение глаз наблюдателя зависит от конкретной обстановки, от полученных им перед просмотром или во время него инструкций, дополнительной информации.

Таким образом, психофизиологические особенности человеческого восприятия характеризуются наличием глубокой обратной связи между органом восприятия (глазом) и органом управления (мозгом) с упорядочиванием найденных объектов по своей значимости. Очевидно, что наиболее адекватное описание данного процесса возможно на основании рекурсивно-иерархических моделей.

В наибольшей степени требованиям снижения информационной избыточности на основе рекурсивно-иерархического представления на сегодняшний день отвечает метод адаптивно-динамической сегментации.

## 6. Адаптивно-динамическая сегментация

Психофизиологическое восприятие изображений человеком включает выделение и анализ объектов и особенностей различного масштаба. При априорно неизвестном масштабе важных для нас особенностей сигнала наиболее эффективным является адаптивно-динамическая сегментация (АДС) сигнала. Это описание легло в основу многих математических моделей, среди которых следует упомянуть динамическую модель дискретного пространства [5], и уже описанные выше методы кратномасштабного и вейвлет-анализа. Окончатель-

ной целью АДС является построение иерархии особенностей сигнала на основе анализа взаимосвязи этих особенностей на различных масштабах. Взаимосвязь между свойствами сигнала на различных масштабах может быть определена либо детерминировано как в случае принудительного разбиения на блоки равного размера при фурье-анализе, или путем половинного дробления области сигнала при вейвлет-анализе, либо адаптивно, например, поиском характерных точек сигнала и объединения их взаимодействий на разных масштабах в древовидную структуру данных [6].

АДС является оригинальным способом адаптивного иерархически структурированного многомасштабного описания сигнала. Главная идея многомасштабного описания с применением АДС заключается в создании некоторого ряда сигналов или изображений, для которого детали на грубых масштабах все более и более подавляются. Проблема выбора характерного масштаба сигнала решается в каждом конкретном случае. Это является результатом того, что в каждом конкретном сигнале и изображении семантически уместные детали существуют в конечном диапазоне масштабов, и необходимо определить эту область масштабов.

В любой задаче иерархической структуризации данных первоочередным вопросом является выбор способа обобщения данных для объединения их в классы более высокого уровня. Для успешного построения иерархической структуры требуется четкая классификация по признакам, имеющим фундаментальное значение для понимания сущности процесса. Под фундаментальным значением в различных случаях понимается неотъемлемая физическая или семантическая характеристика, присущая исследуемому явлению.

Основные требования к методу АДС следующие:

— Постепенное сглаживание (упрощение) сигнала для удаления особенностей на грубых масштабах (Лучше всего, если параметр сглаживания при этом соответствует интуитивному пониманию масштаба);

— Метод нахождения особенностей на каждом уровне сглаживания;

— Причинно-следственная связь, которая устанавливает, что особенности на более грубом уровне сглаживания соотносятся с особенностями на более четких уровнях с меньшей неопределенностью разрешения, однако обратное утверждение может быть и неверным;

— Рекурсивные связи, выражающие соотношения объектов на различных уровнях иерархии для эффективного кодирования разномасштабных объектов.

В зависимости от конкретного приложения, другие желательные свойства подхода АДС могут включать в себя свойства устойчивости, единственности и обратимости представления сигнала.

В основе АДС лежит теория линейного масштабируемого пространства. Согласно данной теории, любой сигнал можно разложить в ряд сигналов с постепенным огрублением детализации и монотонно убывающим числом особенностей, при этом подобное представление будет уникально для каждого сигнала. С технической точки зрения такое разложение эквивалентно операции свертки с фильтрами Гаусса разного временного масштаба. Результат формирования различных уровней детализации для одномерного сигнала показан на рис. 5.

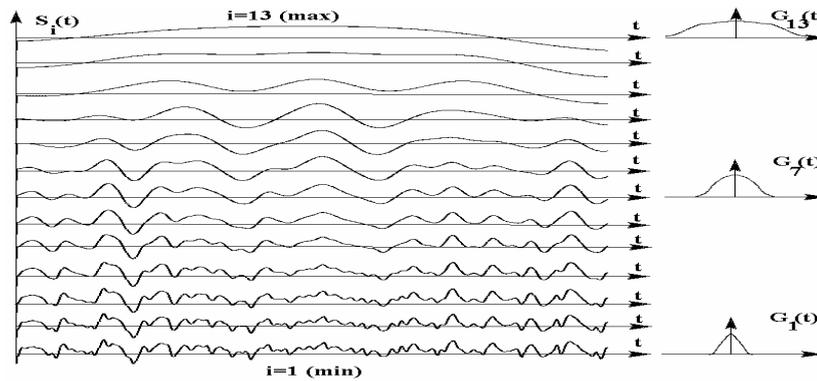


Рис. 5. Разложение сигнала в масштабируемом пространстве

Под особенностями сигнала в данном случае будем понимать участки монотонного поведения между локальными экстремумами. Если взять эти участки в качестве элементов структурного описания, они образуют тернарное дерево, однозначно описывающее исходный сигнал (рис. 6).

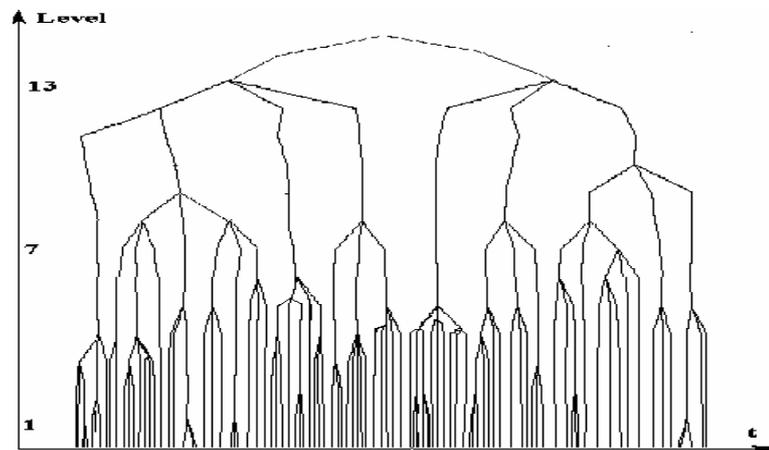


Рис. 6. Древоподобная структура сигнала в методе АДС

В таком древоподобном представлении содержится много узлов, подобных друг другу с точностью до масштабного коэффициента. Таким образом, введение рекурсивных связей в древоподобную структуру позволяет с приемлемой точностью выразить узлы более четкого уровня иерархии через узлы более грубого уровня иерархии (рис. 7).

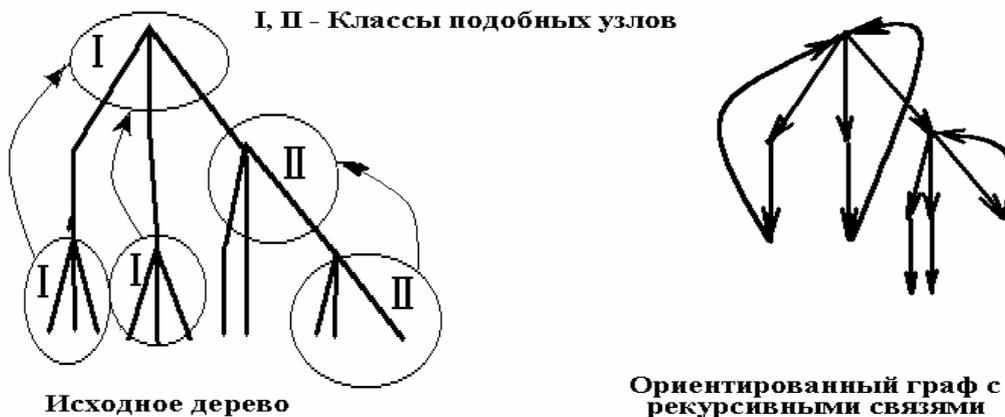


Рис. 7. Введение рекурсивных связей в методе АДС

Как показали исследования, во многих случаях определяемые таким образом элементы древовидной структуры совпадают с идентифицируемыми человеком объектами. Таким образом, АДС по своей сути направлена на устранение психофизиологической избыточности видеоданных.

## **7. Оценка эффективности алгоритмов снижения информационной избыточности, основанных на психофизиологических особенностях человеческого восприятия**

Традиционные методы оценки снижения информационной избыточности, включающие в себя такие критерии как среднеквадратичное отклонение и соотношение сигнал/шум, являются неадекватными в рамках новой парадигмы цифровой технологии. В самом деле, на этапе психофизиологического восприятия изображение представляется человеку не в виде набора отдельных пикселей, оценка изменения цветовых составляющих для которых тривиальна, а в виде иерархически упорядоченной сцены с рекурсивными связями между отдельными ее элементами. Таким образом, в данном случае оценка снижения информационной избыточности является многоуровневой процедурой, включающей в себя:

- оценку информационной насыщенности сцены с точки зрения человеческого восприятия. На данном этапе проводится многоуровневая сегментация с оценкой количества идентифицируемых объектов и глубины их вложенности в иерархической структуре;

- оценку информационной избыточности объектов сцены на различных уровнях иерархии;

- зависимость информационной избыточности объектов сцены от масштаба и уровня иерархии.

Хорошим критерием информационной избыточности, позволяющим получить количественные оценки и основанным на особенностях человеческого восприятия, является стеганографический подход к анализу изображений [2]. В этом случае изображение представляет собой контейнер, в который необходимо вставить произвольное скрытое сообщение без видимых на глаз искажений для человека. Максимально возможный объем такого сообщения и является численной оценкой информационной избыточности данного изображения. Соответственно, изменение данного объема после операций, направленных на снижение информационной избыточности, является оценкой, численно характеризующей величину данного снижения.

Одним из возможных подходов к оценке эффективности алгоритмов снижения информационной избыточности телевизионного сигнала является компонентно-корреляционный анализ. Он заключается в том, что изображения (звукозаписи) на входе и выходе канала подвергаются декомпозиции на составляющие, имеющие различную информационную значимость для зрителя. Раздельно по каждой составляющей вычисляют коэффициент корреляции входного и выходного изображений (до компрессии и после восстановления).

Коэффициенты корреляции по значимым для восприятия составляющим определяют качество передаваемого изображения (чем больше, тем лучше алгоритм передает нужную информацию). Коэффициенты корреляции по незначимым составляющим определяют эффективность алгоритма снижения ин-

формационной избыточности телевизионного сигнала (чем меньше, тем лучше алгоритм исключает ненужные данные). Такой подход позволит охарактеризовать различные алгоритмы в пространстве двух параметров и выбрать оптимальный алгоритм.

Наиболее сложным с методической точки зрения является вопрос создания инструментария для декомпозиции изображения на составляющие, имеющие различную информационную значимость для зрителя. Этот инструментарий должен осуществлять семантико-ориентированный анализ изображения, что является в настоящее время научной проблемой. Поэтому в научном плане основным направлением дальнейших исследований является сравнение методов семантико-ориентированного анализа изображений, в том числе экспериментальное исследование с привлечением испытуемых-добровольцев для выявления метода, наиболее адекватного человеческому восприятию. Предполагается выбрать несколько методов, в том числе вейвлетный и метод адаптивной сегментации, создать набор тестовых изображений с постепенно удаляемой незначимой информацией, получить экспериментальным путем оценки качества изображений испытуемыми и сопоставить эти оценки с параметрами, которые закладывались в алгоритм декомпозиции.

## Заключение

Согласно проведенному в работе обзору, наиболее эффективные алгоритмы снижения информационной избыточности телевизионного сигнала основаны на математических моделях восприятия телевизионного изображения человеком. В свою очередь, особенности человеческого восприятия условно подразделяются на физиологические (early vision в терминах машинного зрения) и психофизиологические.

Несмотря на большое количество опубликованных работ по компрессии телевизионного сигнала, единой теоретической базы для решения данных проблем пока не сформировалось. Большинство предлагаемых алгоритмов получены эвристически, при этом спектр используемых математических инструментов очень широк. Наибольшее распространение и практическое применение на данном этапе получили вейвлет-преобразования, основанные на физиологических особенностях зрительной системы человека. Они позволяют разделить изображение на отдельные компоненты, имеющие различный пространственный масштаб, который, по мнению разработчиков этих алгоритмов, коррелирует с информационной значимостью этих компонентов для зрителя.

Алгоритмы и методы, учитывающие психофизиологические особенности системы зрения человека, в настоящее время еще не получили широкого распространения вследствие сложности объективной оценки их эффективности. Предложенные методы численной оценки их эффективности позволяют проводить их сравнительный анализ, основываясь на интуитивно понятных критериях, что снимает препятствия к их развитию и практическому внедрению.

## Литература

- [1] Александров В. В., Блажис А. К., Скурихин А.В. О новой парадигме информационной среды. // Проблемы информатизации: теоретич. и научно-практич. журнал РАН / Мин. науки и технологий РФ. М., 2000. Вып. 1. С. 52–58.

- [2] Грибунин В. Г., Оков И. Н., Туринцев И. В. Цифровая стеганография. М: СОЛОН-Пресс, 2002. 272 с.
- [3] Бондарко В. М., Данилова М. В., Красильников Н. Н., Леушина Л. И., Невская А. А., Шелепин Ю. Е. Пространственное зрение. СПб.: Наука, 1999. 218 с.
- [4] Добеши И. Десять лекций по вейвлетам. Ижевск: РХД, 2001. 464 с.
- [5] Александров В. В., Арсентьева А. В. Информация и развивающиеся структуры. Л.: ЛНИВЦ АН СССР, 1984. 186 с.
- [6] Скурихин А. В. Рекурсивно-иерархическое представление одномерных фракталоподобных сигналов // Труды СПИИРАН / Под ред. д.т.н. проф. Р.М.Юсупова. СПб.: СПИИРАН, 2003. Вып.1, т. 3. с. 107–117.