

Лекция 2

2 Особенности стандарта MPEG-4 (H.264)

Новым шагом в развитии алгоритмов сжатия изображения стал стандарт MPEG-4. Идея стандарта MPEG-4 заключается не в стандартизации одного продукта, а объединении нескольких подстандартов из которых поставщики могут выбрать один, наиболее соответствующий их задачам.

Стандарт MPEG-4 (H.264) Part 10 или AVC (Advanced Video Coding) — лицензируемый стандарт сжатия видео, предназначенный для достижения высокой степени сжатия видеопотока при сохранении высокого качества. Он содержит ряд новых возможностей, позволяющих значительно повысить эффективность сжатия видео по сравнению с предыдущими (такими, как ASP) стандартами, обеспечивая также большую гибкость применения в разнообразных сетевых средах [5]. По сравнению с MPEG2 (DVD-Video) и MPEG4 ASP (DivX, XviD), сжатие H.264 работает существенно более эффективно, обеспечивая лучшее качество изображения и меньший объём файла.

MPEG4 использует технологию так называемого фрактального сжатия изображений. Фрактальное (контурно-основанное) сжатие подразумевает выделение из изображения контуров и текстур объектов. Контурные представляются в виде т.н. сплайнов (полиномиальных функций) и кодируются опорными точками. Текстуры могут быть представлены в качестве коэффициентов пространственного частотного преобразования (например, дискретного косинусного или вейвлет-преобразования). Диапазон скоростей передачи данных, который поддерживает формат сжатия видео изображений MPEG 4, гораздо шире, чем в MPEG 2. Дальнейшие разработки специалистов направлены на полную замену методов обработки, используемых форматом MPEG 2. Формат сжатия видео изображений MPEG4 поддерживает широкий набор стандартов и значений скорости передачи данных. MPEG 4 включает в себя методы прогрессивного и

чересстрочного сканирования и поддерживает произвольные значения пространственного разрешения и скорости передачи данных в диапазоне от 5 кбит/с до 10 Мбит/с. В MPEG 4 усовершенствован алгоритм сжатия, качество и эффективность которого повышены при всех поддерживаемых значениях скорости передачи данных.

2.1 Возможности стандарта H.264

1) Многокадровое предсказание. Использование сжатых ранее кадров в качестве опорных (то есть с заимствованием части материала из них) куда более гибко, чем в предыдущих стандартах. Позволяется использование до 32 ссылок на другие кадры, тогда как в ASP и более ранних число ссылок ограничено одним или, в случае В-кадров, двумя кадрами. Это поднимает эффективность кодирования, так как позволяет кодеру выбирать для компенсации движения между большим количеством изображений. В большинстве сцен данная функция обеспечивает не очень большое улучшение в качестве и не даёт заметного понижения битрейта. Однако, для некоторых сцен, например, с частыми повторяющимися участками, возвратно-поступательным движением и т. п. данный подход при сохранении качества позволяет очень сильно снизить затраты битрейта [6].

2) Независимость порядка воспроизведения изображений и порядка опорных изображений. В предшествующих стандартах устанавливалась жёсткая зависимость между порядком следования изображений для использования при компенсации движения и порядком следования изображений при воспроизведении. В новом стандарте эти ограничения в значительной мере устранены, что позволяет кодеру выбирать порядок изображений для компенсации движения и для воспроизведения с высокой степенью гибкости, которая ограничена только объёмом памяти, который гарантирует возможность декодирования. Устранение ограничения также позволяет в ряде случаев устранить дополнительную задержку, ранее связанную с двунаправленным предсказанием.

3) Независимость методов обработки изображений и возможности их использования для предсказания движения. В предшествующих стандартах изображения, закодированные с использованием некоторых методов (например, двунаправленного предсказания), не могли использоваться в качестве опорных для предсказания движения других изображений видеопоследовательности. Устраняя это ограничение, новый стандарт обеспечивает кодеру большую гибкость и, во многих случаях, возможность использовать для предсказания движения изображение, более близкое по содержанию к кодируемому.

4) Точность до четверти пикселя при компенсации движения обеспечивает очень высокую точность описания движущихся областей (что особенно актуально для медленного движения). Цветность, как правило, хранится с разрешением, уменьшенным вдвое по вертикали и горизонтали (прореживание цвета), поэтому компенсация движения для компонента цветности использует точность в одну восьмую пикселя цветности.

5) Квантование – логарифмическое управление длиной шага для упрощения распределения битрейта кодером и упрощенного вычисления обратной длины квантования.

6) Частотно-оптимизированные матрицы масштабирования квантования, выбираемые кодером для оптимизации квантования на основе человеческих особенностей восприятия (поддерживается не во всех профилях).

7) Разбиение данных — функция, обеспечивающая разделение данных разной важности (например, векторы движения и другая информация предсказания имеет большую значимость для представления видеоконтента) по разным пакетам данных с разными уровнями защиты от ошибок (поддерживается не во всех профилях).

8) Избыточные части. Возможность посылки кодером избыточного представления областей изображений, позволяя воспроизвести области

изображений (обычно с некоторой потерей качества), данные о которых были потеряны в процессе передачи (поддерживается не во всех профилях).

9) Нумерация кадров, позволяющая создание «подпоследовательностей» (включая временное масштабирование включением дополнительных кадров между другими) а также обнаружение (и скрывание) потерь целых кадров при сбоях канала или пропаже пакетов [1].

2.2 Преимущества и основные недостатки стандарта H.264

По сравнению с MPEG2 (DVD-Video) и MPEG4 ASP (DivX, XviD), сжатие H.264 работает существенно более эффективно, обеспечивая лучшее качество изображения (вплоть до недостижимого для MPEG2 и MPEG4 ASP уровня) и меньший объём файла.

Главным недостатком H.264 являются заметно более высокие требования к оборудованию для кодирования и воспроизведения видеофайлов же оборудовании во многих случаях воспроизводится довольно таки хорошо [6].

Формат запатентован, и создатели кодеков обязаны платить за их распространение путём покупки лицензий. С 2011 года MPEG LA могла бы начать взимать плату и с тех, кто участвует в кодировании и/или бесплатном предоставлении пользователям видеопотока в AVC. Однако позже этот срок был изменён на 2015 год, а 26 августа 2010 года компания MPEG LA объявила, что за бесплатное предоставление пользователям видеопотока в H.264 плата взиматься не будет.

2.3 Поддержка стандарта H.264

H.264 принят в качестве стандарта для сжатия видео высокой чёткости (HD, HDTV), распространяемого на оптических носителях нового поколения — Blu-ray и HD DVD, используется в мобильных устройствах, поддерживается в Apple QuickTime, получает распространение в системах

цифрового телевидения, видеоконференцсвязи, видеонаблюдения и проч. Adobe Flash Player, являющийся стандартом де-факто для мультимедийных web-приложений и онлайн-видеохостингов вроде YouTube, поддерживает прямое воспроизведение H.264-видео начиная с версии 9.0.115, вышедшей в конце 2007 г. [6].

Для воспроизведения видеофайлов формата H.264 достаточно установить следующие бесплатные программные продукты: Haali Media Splitter — для чтения данных из контейнеров MKV и MP4; ffdshow tryouts — универсальный декодер с поддержкой H.264; AC3Filter — для декодирования звукового потока AC3; программный плеер Media Player Classic, VLC media player или какой-либо др.

2.4 Принципы помехоустойчивого канального кодирования сигналов изображения и звука

2.4.1 Порядок операций по защите от ошибок в передающей и приемных частях систем цифрового телевидения

В современных системах цифрового телевидения применяется каскадное кодирование транспортного цифрового потока MPEG2, поскольку вероятность ошибок в используемых каналах связи часто может быть настолько велика, что одна ступень кодирования не обеспечивает необходимой помехоустойчивости.

Операции по защите от ошибок на передающей стороне выполняются в следующем порядке:

- скремблирование;
- внешнее кодирование с использованием кода Рида – Соломона;
- внешнее перемежение;
- внутреннее кодирование (ближе к каналу связи)– сверточное кодирование.

В приемной части системы операции выполняются в обратном порядке:

- внутреннее декодирование сверточного кода (алгоритм Витерби),
- внешнее дегермежение,
- декодирование Рида – Соломона;
- дескремблирование.

Ниже рассматриваются особенности и важнейшие характеристики перечисленных операций по защите цифровых телевизионных сигналов от ошибок.

2.4.2 Принципы скремблирования

Для передачи цифровых данных на большие расстояния используется последовательная передача, когда все биты отсчета передаются по одной физической линии методом временного уплотнения (мультиплексирования). Разумеется, при этом значительно возрастают тактовая частота и затухание в канале, однако современная техника успешно справляется с такими трудностями. Однако все операции обработки цифровых сигналов, аналого-цифрового и цифро-аналогового преобразования производятся в параллельной форме. Переход от последовательной к параллельной передаче и обратно производится с помощью соответствующих преобразователей в параллельный (SPC – Serial to Parallel Converter) или последовательный (PSC – Parallel to Serial Converter) код. В процессе передачи приемник должен точно определить, к какому временному интервалу отнести тот или иной принятый бит, и это задача тактовой синхронизации. Возможны несколько путей решения указанной задачи. В цифровом вещании существенное повышение точности синхронизации обеспечивается за счет канального кодирования, обеспечивающего отсутствие постоянной составляющей, и скремблирования (рандомизации) цифрового потока. При этом тактовая частота отдельно не передается, а выделяется на приеме из полезных данных.

В наиболее распространенном последовательном интерфейсе SDI (Serial Digital Interface - последовательный цифровой интерфейс), описанном в Рекомендации ВТ.656 и стандарте SMPTE-259М, для повышения

надежности выделения тактовой частоты используется инвертированный NRZ (Non Return to Zero) код и сверточное скремблирование. В качестве исходного используется рассмотренный выше цифровой компонентный сигнал "4:2:2" в параллельном формате.

Входной цифровой поток имеет фиксированную длину пакетов. Суммарная длина пакета MPEG-2 транспортного мультиплексора 188 байт, который включает в себя 1 байт синхронизации. Порядок обработки на передающей стороне всегда начинается со старшего разряда синхронизирующего байта (см. рис.2.9). Каждый пакет транспортного потока начинается с синхробайта 0x47. Эти синхробайты не скремблируются. Генерация ПСП повторяется после скремблирования каждых 8 пакетов транспортного потока. Чтобы обеспечить синхронизацию дескремблера в приемнике (цикловую синхронизацию), синхробайт первого пакета в группе из 8 пакетов инвертируют с 47_{HEX} в $B8_{HEX}$.

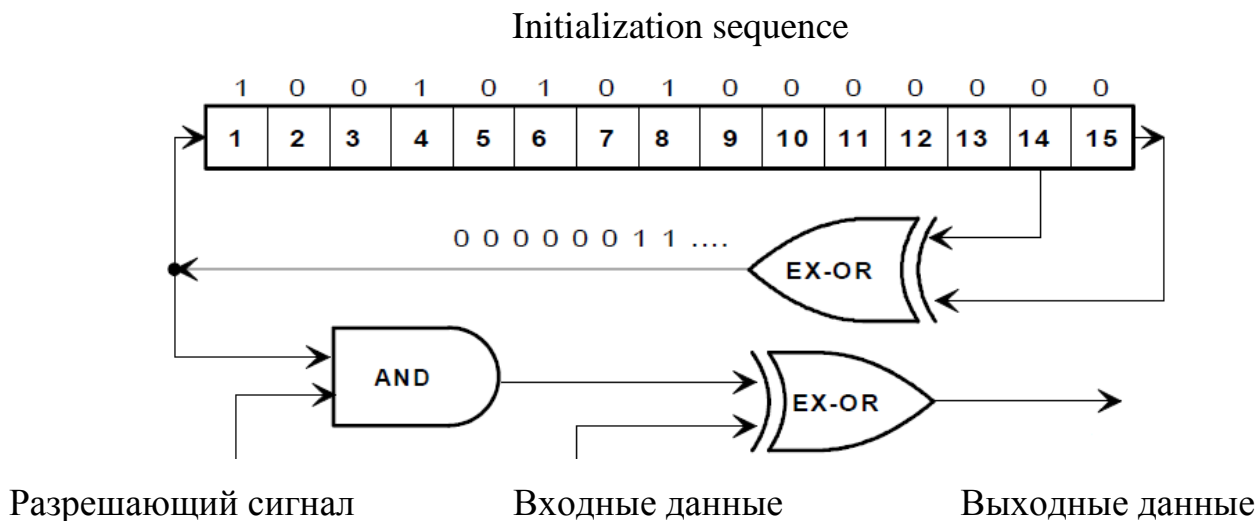


Рисунок 2.1 – Принципиальная схема рандомизатора/дерандомизатора

Первый бит с выхода генератора псевдослучайной двоичной последовательности должен быть применен к первому биту

инвертированного MPEG-2 синхронизирующего байта ($B8_{HEX}$). Для поддержки других функций синхронизации, генератор псевдослучайной последовательности постоянно функционирует, но его выход заблокирован, оставляя байты нерандомизированными. Период псевдослучайной последовательности должен быть 1503 байта.

В результате скремблирования характеристики цифрового сигнала становятся близкими к характеристикам случайного сигнала: устраняются длинные серии нулей и единиц, количества этих символов примерно выравниваются, а переходы между ними происходят в случайном порядке. Это обеспечивает более точное восстановление тактовой частоты. Кроме того, выравнивается энергетический спектр сигнала, что также способствует повышению помехоустойчивости. Еще одна функция скремблирования - шифровка сигнала для ограничения доступа к ТВ программам.

В приемнике выполняется обратная операция - дескремблирование. Она заключается в сложении скремблированного сигнала по модулю 2 с той же самой псевдослучайной последовательностью.

2.4 Принципы помехоустойчивого кодирования

Одним из мощных средств борьбы с ошибками являются корректирующие коды. Процесс помехоустойчивого кодирования состоит в том, что наборы из k информационных символов отображаются в кодовые последовательности, состоящие из n символов, где $n > k$. Наличие избыточных символов (их еще называют проверочными) позволяет отслеживать появление ошибок и при обнаружении исправлять их. Отношение $R = k/n$ называется *относительной скоростью кода*. Естественно, что при постоянной информационной скорости введение избыточности при кодировании повышает скорость модуляции и расширяет полосу используемых частот.

Для оценки эффективности помехоустойчивого кодирования сравнивают отношение энергии E_b , приходящейся на один бит, к

спектральной плотности мощности шума N_0 в системе с кодированием и в базовой системе без кодирования и определяют разницу в значениях E_b/N_0 при заданной вероятности ошибки (рис. 2.10). Эта разница, измеряемая в децибелах и называемая *энергетическим выигрышем кода* (ЭВК), может быть использована для сравнения разных кодов.



Рисунок 2.2 – Оценка энергетического выигрыша кода (ЭВК)

Для повышения эффективности кодирования в последние годы в технике связи широко применяют каскадные коды - последовательное кодирование символов двумя различными кодами, как показано на рис. 1.16.



Рисунок 2.3 – Организация каскадного канального кодирования/декодирования

Коды называются *внутренним* и *внешним* в зависимости от того, какое положение по отношению к каналу передачи они занимают. Внешний код - обычно код Рида-Соломона, в качестве внутреннего могут использоваться различные коды- сверточные, короткие блочные, биортогональный код и др.

В табл.2.1 приведен результирующий выигрыш кодирования для разных комбинаций внутреннего и внешнего кодов при двух значениях вероятности ошибки - 10^{-5} и 10^{-8} .

Таблица 2.1 – Энергетический выигрыш для разных вариантов кодирования

Вариант кодирования	ЭВК	
	$P_{\text{ош}}=10^{-5}$	$P_{\text{ош}}=10^{-8}$
Рида-Соломона + Витерби	6,5...7,5	8,5...9,5
Рида-Соломона + биортогональный	5...7	7...9
Рида-Соломона + короткий блочный	4,5...5,5	6,5...7,5
Витерби	4...5,5	5...6,5
Блочный код (жесткое решение)	3...4	4,5...5,5
Сверточный код (пороговое декодирование)	1,5...3	2,5...4,0

Для оценки обнаруживающей и исправляющей способности кода используют понятие кодового расстояния $d_{\text{мин}}$ - наименьшего из расстояний между любой парой кодовых последовательностей. Расстоянием между двумя последовательностями (его еще называют расстоянием Хэмминга) называется число позиций, в которых они отличаются друг от друга. Для вычисления расстояния между двоичными последовательностями используют посимвольное сложение по модулю 2. При декодировании ошибка обнаруживается, если ее кратность (число ошибочных символов на длине блока) не превышает $d_{\text{мин}}-1$. Ошибка будет исправлена при $t \leq [(d_{\text{мин}}-1)/2]$, где квадратные скобки означают целую часть числа. Декодер, который декодирует каждую принятую последовательность в ближайшее к ней по расстоянию Хэмминга кодовое слово, выбирает то кодовое слово, условная вероятность передачи которого максимальна, и потому называется декодером максимального правдоподобия.

Современная теория кодов достаточно развита и содержит детальную классификацию. Все применяемые коды можно разбить на две большие группы: *блочные*, в которых кодирование и декодирование производится в

пределах определенного участка кодовой последовательности - блока, и *древовидные*, в которых обработка символов производится непрерывно, без разделения на блоки. Часть кодов относится к разряду *линейных*, в которых кодовые последовательности представлены как элементы линейного векторного пространства. Наконец, можно применить разбиение на *коды, исправляющие независимые случайные ошибки*, и *коды, исправляющие пакетные ошибки*. Несмотря на появление мощных кодов, исправляющих пакетные ошибки, часто оказывается более выгодным использовать *коды, исправляющие независимые случайные ошибки, совместно с устройствами перемежения и восстановления*.

Линейные древовидные коды называются *сверточными*, поскольку процесс кодирования можно представить как специальную линейную операцию - свертку входной последовательности с импульсным откликом кодера. Сверточные коды при вычислении выходного символа учитывают текущий входной символ и несколько предшествующих. Общее число учитываемых символов K называется кодовым ограничением.

2.5 Внешнее кодирование с использованием кодов Рида-Соломона

Среди большого числа блоковых кодов наибольший интерес представляют коды Рида-Соломона - подкласс не двоичных кодов размерности $q=2^m$, m - целое, обеспечивающих наибольшее минимальное расстояние между словами при заданной длине блока, равное теоретическому пределу. Их используют для исправления пакетных ошибок (в сочетании с перемежением), а также в каскадных системах кодирования в качестве внешних кодов.

Способность кода Рида-Соломона корректировать несколько ошибок обусловлена тем обстоятельством, что к исходным данным добавляется несколько избыточных символов и принятое кодовое слово делится не на один, а на несколько многочленов, давая соответствующее число синдромов. Пусть, например, добавлены два символа. Тогда одновременное решение

двух получившихся уравнений дает значения двух неизвестных, одна из которых определяет положение ошибочного символа в кодовом слове и называется локатором, а другая дает структуру ошибки и именуется корректором. Для восьмибитовых символов полная длина кодового слова должна составлять $2^8 - 1 = 255$ символов, из них 239 информационных и 16 проверочных. Для уменьшения объема расчетов и других целей код можно укоротить, заменив часть кодовых слов нулевыми кодовыми комбинациями на передаче и исключая соответствующие символы на приеме.

Первая ступень кодирования (*внешнее кодирование*) с использованием кодов *Рида-Соломона* выполняется сразу же после скремблирования во всех видах цифрового ТВ вещания. В каждый пакет транспортного потока вводится 16 проверочных байтов, так что размер пакета увеличивается до $188+16 = 204$ байтов. Проверочные байты располагаются в конце пакета. Значения проверочных байтов вычисляются по значениям всех 188 информационных байтов пакета.

В декодере для каждого принятого транспортного пакета по содержащимся в нем 204 байтам вычисляют положения ошибочных байтов и значения ошибок в них. Ошибки корректируются, если это возможно. Если же коррекция невозможна (например, ошибочных байтов слишком много) данные в пакете не изменяются, а сам пакет помечается путем установки флага (первый бит после синхробайта), как содержащий неустранимые ошибки. В обоих случаях 16 проверочных байтов удаляются, и после декодирования длина транспортного пакета становится равной 188 байтов.

Основными параметрами применяемого кода Рида – Соломона являются относительная скорость кодирования R и исправляющая способность кода t .

Относительная скорость кодирования составляет

$$R = 188 / 204 = 0,92.$$

Исправляющая способность кода обеспечивает коррекцию до 8 байтов с ошибками в каждом пакете транспортного потока, т.е. исправляющая способность $t = 8$ байтов.

2.6 Перемежение и деперемежение

Эффективным методом уменьшения влияния пакетных ошибок является внешнее *перемежение* или перемешивание (англ. - interleaving). Данные перед передачей по каналу связи переставляются в заданном порядке, а в приемной части восстанавливается исходный порядок, то есть выполняется *деперемежение*. При этом пакетная ошибка, возникшая в канале связи, превращается в набор рассредоточенных во времени одиночных ошибок, которые проще обнаруживаются и исправляются с помощью кодов, исправляющих ошибки.

Принцип выполнения перемежения (внешнего) и деперемежения поясняется простым примером на рис.7. Исходный цифровой сигнал представляет собой последовательность 4-разрядных двоичных слов, передаваемых бит за битом (рис.2.4,а). Перемежение выполняется в пределах каждых 4 слов, то есть в пределах отрезка цифрового сигнала, содержащего 16 бит. Числа показывают номер бита в этом отрезке. В результате перемежения биты переставляются (рис.2.4,б). Биты, искаженные действием пакетной ошибки, отмечены звездочками. В результате деперемежения (рис.2.4,в) восстанавливается исходный порядок битов и искаженные биты рассредоточиваются и обеспечиваются условия для эффективного исправления локальных ошибок декодером Рида-Соломона.

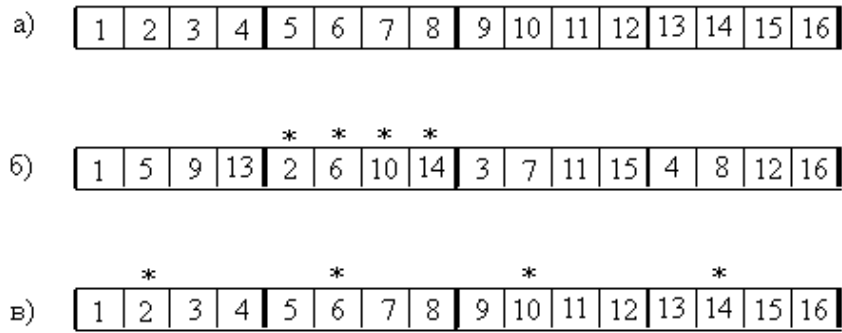


Рисунок 2.4 – Внешнее перемешивание и деперемешивание

2.7 Внутреннее сверточное кодирование

Сверточные коды используются при низком отношении сигнал-шум, когда исправляющей способности блочных кодов при разумной длине блока оказывается недостаточно. Внутреннее сверточное кодирование осуществляется без разделения передаваемого сообщения на блоки или кодовые слова. Последовательность информационных символов непрерывно поступает на вход кодера, на выходе которого формируется последовательность передаваемых символов. Сверточные коды получили широкое применение, т.к. оказались во многих случаях лучше блочных при сопоставимой сложности аппаратной реализации кодера и декодера.

Параметры сверточного кодера: память кодера k – число разрядов сдвигового регистра, кодовая скорость k/n – отношение количества символов во входной и выходной последовательностях, кодовое ограничение – число символов входной последовательности, влияющих на символ выходной последовательности. Пример кодера с

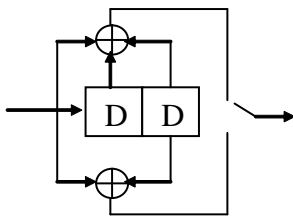


Рисунок 2.5 $k/n = 1/2$, приведен на рис.2.5. Биты входной информационной последовательности вводятся по одному в сдвиговый регистр. За один такт коммутатор снимает с сумматоров и выдает на выход кодера 2 бита. Каждый бит выходной последовательности зависит от трех битов входной последовательности. Кодовая скорость может быть повышена

вычеркиванием некоторых символов на выходе кодера (перфорация, выкалывание кода) по определенному алгоритму.

Важными параметрами сверточного кода являются скорость кода R , кодовое ограничение K и свободное расстояние кода.

Скорость кода R – это отношение количества входных информационных битов обрабатываемых кодером одновременно к количеству выходных битов с кодера. Она зависит от схемы кодера, т.е. от количества входных информационных битов (одновременно хранящихся в памяти кодера) по которым строится выходной код. Например, $R=1/2$ означает, что на один входной информационный бит кодер выдает два выходных бита. Скорость кода может быть равна также $2/3$, $3/4$, $5/6$, $7/8$.

Кодовое ограничение K – число символов входной информационной последовательности, запоминаемых в кодере и используемых для получения выходных символов. Например, в схеме на рис.2.5 кодовое ограничение $K=3$.

Свободное расстояние кода – аналог минимального расстояния Хэмминга между разрешенными кодовыми словами блочного кода. Свободное расстояние кода – количество единиц на выходе кодера при входной информационной последовательности, состоящей из одной единицы и всех остальных нулей. Например, если на входе кодера информационная последовательность 0 1 0 0 0 0 0, то на выходе кодера при скорости кода $R = 1/2$ получим 00 11 01 11 00 00 00. Количество единиц стало равно 5, значит свободное расстояние кода равно 5.

Удобным средством описания работы кодера является решетчатая или сетевая диаграмма.

Решетчатая диаграмма (trellis diagram) является удобным представлением процесса изменения состояний сверточного кодера при поступлении на его вход битовой последовательности. Пример приведен для кодера с четырьмя состояниями (см. рис.2.6). Пунктирные линии показывают изменения состояний при поступлении на вход кодера нуля, сплошные линии – при

поступлении единицы. Каждый переход помечен соответствующей кодовой комбинацией, появляющейся на выходе кодера.

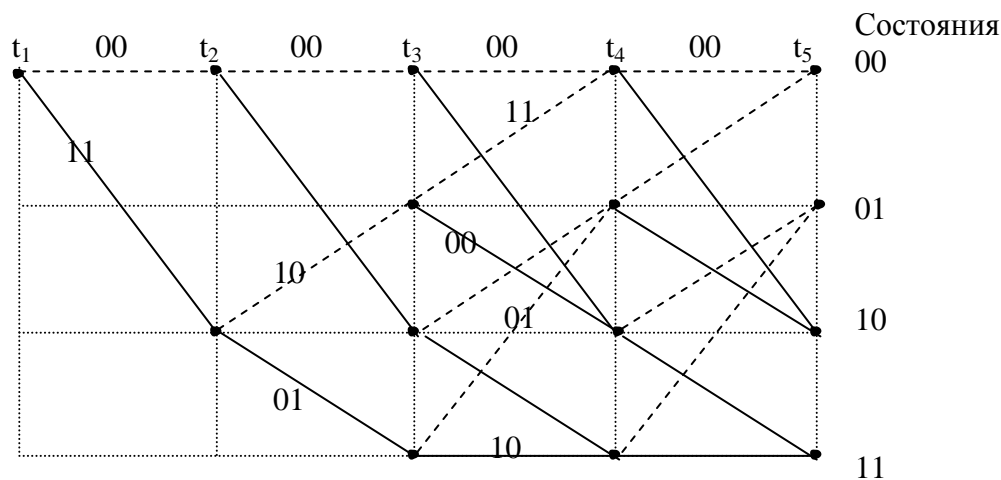


Рисунок 2.6 – Решетчатая диаграмма

Каждой входной последовательности соответствует определенный путь в решетке и соответствующая этому пути «правильная» выходная последовательность. Начальное состояние - 00.

Один из алгоритмов декодирования (алгоритм Витерби, или алгоритм максимального правдоподобия), сравнивает принятую декодером кодовую комбинацию с правильными, соответствующими различным путям в решетке, и заменяет на ту, которая по кодовому расстоянию ближе всех к принятой. Число различных путей в решетке очень велико. Объем вычислений минимизируется благодаря использованию итерационного (пошагового) алгоритма, аналогичного алгоритму динамического программирования.

Как видно из диаграммы, после K переходов ($K=3$) картина переходов повторяется. В каждое состояние ведут 2 пути. Для каждого пути длиной в i переходов, начинающегося из начального состояния и заканчивающегося в момент t_i (на первом шаге вычислений $i=3$), определяется метрика - кодовое расстояние между принятой последовательностью и правильной последовательностью, соответствующей данному пути. Из всех путей, ведущих в данное состояние, выбирается один путь с минимальным кодовым

расстоянием, остальные пути отбрасываются как "тупиковые". После просмотра путей длиной в i переходов просматриваются и отбрасываются тупиковые пути длиной в $i + 1$ переходов и т.д. Начиная с некоторого момента t_i , начальный участок у всех "выживших" путей оказывается одним и тем же. Соответствующая этому участку кодовая комбинация считается достоверной. По мере роста номера i длина общего участка, а следовательно и скорректированной части принятой кодовой последовательности, увеличивается.

Для декодирования сверточного кода применяется алгоритм Витерби. Он позволяет из множества возможных путей, приводящих к декодируемому символу, выбрать относительно небольшое число путей, являющихся наиболее правдоподобными, и определить правильное значение символа.

В цифровом телевизионном вещании, как правило, используется сочетание сверточного кодирования с декодированием по алгоритму Витерби, хорошо работающего при низких отношениях сигнал-шум, и кода Рида-Соломона. Для канала с аддитивным белым шумом, необходимое для передачи отношение E_b/N_0 , определяется неравенством

$$\frac{E_b}{N_0} \geq \frac{1}{2R} (2^{2R} - 1) . \quad (2.1)$$

Например, при $R=1/2$ такое сочетание обеспечивает суммарную вероятность ошибки 10^{-10} при отношении энергии бита к спектральной плотности шума $E_b/N_0 = 2$ или 3 дБ.

При передаче по каналам с ограниченной полосой (типовой режим для вещательных каналов) метод кодирования для достижения максимальной эффективности должен учитывать также и выбранную схему модуляции.

2.8 Модуляция в системах цифрового телевизионного вещания

Хотя модулированные сигналы являются аналоговыми, однако рассмотрение процесса модуляции весьма важно, так как от его удачной реализации зависят многие характеристики системы цифрового телевизионного вещания.

В системах цифрового вещания широко применяются амплитудная и фазовая модуляция (манипуляция). При амплитудной манипуляции несущей бинарным сигналом на выходе модулятора получается колебание несущей частоты

$$s(t) = S(t) \cos[\omega_0 t - \varphi(t)] \quad , \quad (2.2)$$

амплитуда которого $S(t)$ принимает одно из двух значений S_0 или S_1 , соответствующих 0 и 1. Многоуровневая, или многопозиционная АМ с кратностью M формирует сигнал с M возможными значениями амплитуды, при этом начальная фаза сигнала $\varphi(t)=\varphi_0$ и в процессе модуляции не изменяется. Спектр амплитудно манипулированного сигнала содержит несущую и две боковые полосы, повторяющие форму спектра модулирующего сигнала. В некоторых случаях используют сигналы с полным или частичным подавлением несущей, одной из боковых полос или несущей и боковой полосы одновременно.

При фазовой модуляции (манипуляции) амплитуда сигнала остается постоянной, а начальная фаза при манипуляции скачкообразно изменяется на величину, зависящую от кратности модуляции M . Дискретные значения фазы сигнала $\varphi_i=2\pi a_i/M$, где a_i - M -ичный символ из множества $\{0,1,2,\dots,M-1\}$. Обычно $M=2^k$, $k=1,2,3\dots$. Тогда каждый символ a_i соответствует набору из $K=\log_2 M$ двоичных информационных символов, а его длина $T_c = kT$, где T - длительность двоичного символа. При $M = 2$ получаем сигналы

двухпозиционной (BPSK), $M = 4$ - четырехпозиционной (QPSK), $M=8$ - восьмипозиционной (8-PSK) модуляции.

Модулированный сигнал (2.6) можно представить в виде двух квадратурных (со сдвигом на 90°) составляющих

$$s(t) = [S(t) \cos \varphi(t)] \cos \omega_0 t + [S(t) \sin \varphi(t)] \sin \omega_0 t \quad . \quad (2.3)$$

Если в соответствии с (2.7) промодулировать, например, бинарными сигналами две несущие $\cos \omega_0 t$ и $\sin \omega_0 t$ с фазовым сдвигом 90° и полученные сигналы сложить, результирующий сигнал будет уже четырехпозиционным, и такая модуляция называется квадратурной, а две компоненты сигнала - синфазной (I - inphase) и квадратурной (Q - quadrature). Обычно полагают, что вектор I совпадает с осью абсцисс, а вектор Q ориентирован вдоль оси ординат. Если сигнальные векторы дискретно модулируются не только по фазе, но и по амплитуде получается - квадратурная АМ (QAM).

Представление сигналов в виде суммы квадратурных составляющих подсказывает простой способ их формирования в модуляторе. В частности, сигналы QPSK можно получить как сумму двух сигналов BPSK, несущие колебания которых отличаются на 90° . Схема такого модулятора показана на рис. 2.6 Модулирующие последовательности в квадратурных каналах получают расщеплением исходной информационной последовательности на четные и нечетные символы. Длительность символа в каждом канале равна $T_c = 2T$. Сигналы BPSK квадратурных каналов складываются, образуя четырехфазный сигнал QPSK.

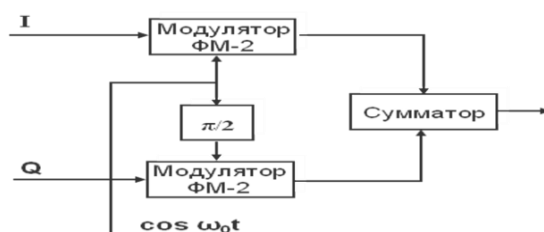


Рисунок 2.6 – Структурная схема квадратурного модулятора

При одновременной смене символов в каждом из квадратурных каналов модулятора в сигнале QPSK происходит скачок фазы на π , что вызовет на выходе фильтра провал огибающей до нуля. Исключить скачки фазы на π возможно, если разнести во времени моменты смены фаз в квадратурных каналах. Обычно модулирующие сигналы в одном из каналов смещают на величину $T_c/2$ (так, называемая смещенная PSK).

От формы модулирующих сигналов в каналах зависит наличие или отсутствие межсимвольных помех и общая ширина полосы частот, занимаемая PSK сигналом. Обычные прямоугольные сигналы в этом смысле далеко не оптимальны. Существенного сужения спектра удастся добиться скруглением элементарных модулирующих сигналов. Наиболее часто применяют спектры Найквиста с так называемым косинусным скруглением, которое задается соотношением:

$$1, \quad 0 \leq \Omega \leq \pi(1-\alpha)/T,$$

$$S(\Omega) = \left\{ \cos^2 \left[\frac{T}{4\alpha} \left(\Omega - \frac{\pi(1-\alpha)}{T} \right) \right], \frac{\pi(1-\alpha)}{T} \leq \Omega \leq \frac{\pi(1+\alpha)}{T} \right\}, \quad (2.4)$$

$$0, \quad \Omega > \pi(1+\alpha)/T.$$

Коэффициент α характеризует степень скругления спектра. На рис. 2.7 приведены частные случаи спектра Найквиста. Условием отсутствия межсимвольных помех является передача символов со скоростью $1/T$.

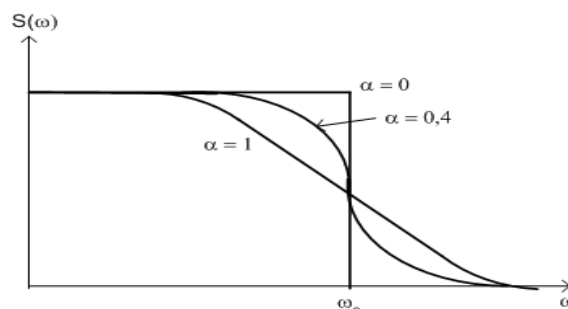


Рисунок 2.7 – Необходимая АЧХ канала при различных коэффициентах α

Сравнение свойств различных сигналов удобно производить, используя их геометрическое представление. Многомерные сигналы можно изобразить в виде точек (концов векторов) на плоскости. На рис.2.8 показаны типичные ансамбли многопозиционных сигналов с квадратурными фазовой и амплитудной манипуляцией. Пространство сигналов модулированной несущей представляет собой дискретные положения вершин суммарного вектора $(I+Q)$ в системе координат I, Q . В случае QPSK вектора I и Q имеют одинаковые и постоянные амплитуды, дискретно меняются только фазы, которые могут принимать значения 0° или 180° - для вектора I и 90° и 270° - для вектора Q . При этом энергия PSK сигналов в процессе модуляции не меняется, и вершины суммарного вектора $I+Q$ при переходе от одного фазового положения до другого описывают окружность. В случае QAM концы сигнальных векторов располагаются в узлах прямоугольной решетки. Видно, что при $M = 4$ вектора QPSK и 4-QAM совпадают.

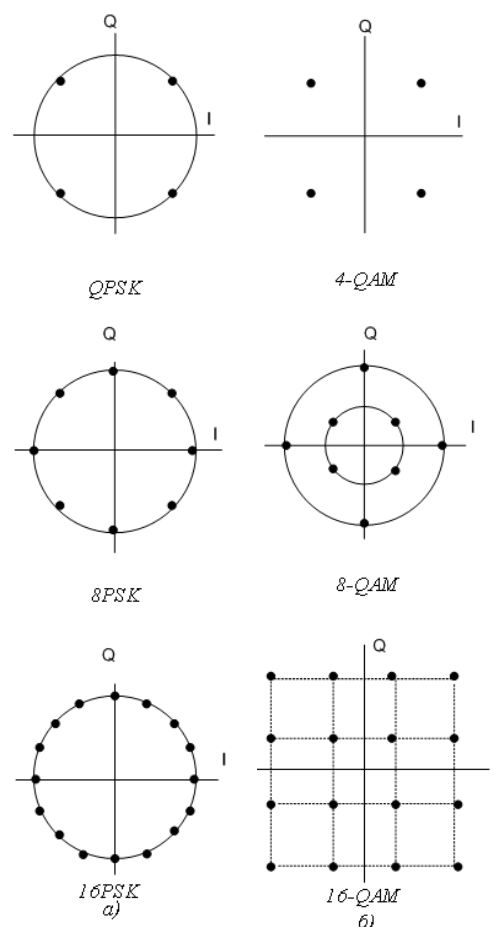


Рисунок 2.8 – Созвездия PSK и QAM сигналов при различных M

В ансамблях многопозиционной QAM дискретизация по амплитуде возможна с равномерным и неравномерным шагом. В первом случае мы имеем дело с *неиерархической модуляцией*, обеспечивающей одинаковую помехоустойчивость для всех передаваемых комбинаций, во втором случае, с *иерархической модуляцией*, когда некоторые компоненты потока могут передаваться с повышенной помехоустойчивостью. С точки зрения помехоустойчивости важно сохранять достаточно большим минимальное расстояние между двумя соседними точками в ансамбле. Доказано, что это условие выполняется при размещении точек в узлах квадратной решетки.

В многопозиционных ансамблях каждый сигнал содержит информацию о k двоичных информационных символах ($M = 2^k$). Для минимизации вероятности ошибки на символ необходимо так сопоставить каждому сигналу набор символов, чтобы последовательности, соответствующие соседним сигналам, отличались наименьшим числом двоичных символов. Этим свойством обладают некоторые манипуляционные коды, в частности, широко применяемый код Грея. На рис.2.9 приведены примеры построения кодовых комбинаций манипуляционного кода Грея.

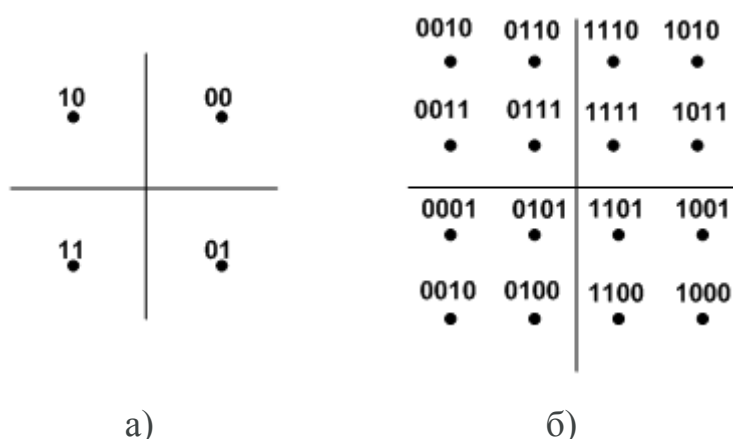


Рисунок 2.9 – Примеры построения кодовых комбинаций манипуляционного кода Грея

Выше мы говорили о построении кода, учитывающего вид модуляции и расстояния между сигналами в сигнальном пространстве. Оказывается, что

при многомерных сигналах свойства кода сами по себе не обязательно определяют вероятность ошибки. Важно, какие евклидовы расстояния получаются при сочетании способа кодирования и схемы модуляции. Напомним, что евклидовым расстоянием между двумя точками a и b n -мерного пространства называется величина

$$d = \sqrt{\sum_{i=1}^n (a_i - b_i)^2}. \quad (2.5)$$

Основным параметром, характеризующим в случае сверточных кодов комбинацию варианта кодирования и способа модуляции, является свободное евклидово расстояние $d_{св}$, определяемое как минимальное евклидово расстояние между любыми парами кодовых последовательностей v_k и v_j . Если необходимо согласовать выбор кода и выбор сигналов, то критерием в этом случае служит не вероятность ошибки двоичного символа, а асимптотический выигрыш от кодирования, который определяется формулой

$$G_a = 20 \lg(d_{св} / d_{баз}), \quad (2.6)$$

где $d_{баз}$ - минимальное евклидово расстояние между различными последовательностями в базовой системе без кодирования.

В качестве примера рассмотрим кодирование для 8-QAM с использованием сверточного кода и декодирования Витерби при двух информационных битах на передаваемый символ. Требуемые для 8-QAM три бита на символ составим следующим образом: первые два бита представляют собой выходную последовательность сверточного кода с $R = 1/2$, а в качестве третьего бита используется незакодированный второй бит исходных данных. Сигнальная диаграмма такого кода приведена на рис.2.10, символы сверточного кода подчеркнуты. Максимально разделенными при таком отображении оказываются пары сигнальных точек, в которых закодированные биты совпадают, а незакодированные информационные

биты различны. Базовая система с двумя информационными битами на символ, с которой нужно сравнивать данную систему - 4-QAM. Асимптотический выигрыш описанной системы зависит от кодового ограничения K. При K = 7 он достигает 4 дБ. Этот код используется в американском стандарте цифрового вещания ATSC под названием трелисный (решетчатый) код.

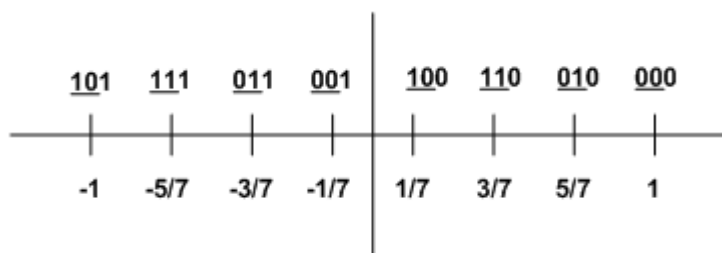


Рисунок 2.10 – Сигнальная конструкция трелисного кода 8T-VSB

Сравнение различных методов модуляции проводят по нескольким критериям: помехоустойчивости, скорости передачи на единицу полосы, степени влияния искажений в тракте передачи и т.д.

Помехоустойчивость определяется зависимостью вероятности ошибки от отношения энергии бита E_b к спектральной плотности шума N_0 и для сигнала BPSK рассчитывается по формуле:

$$p = F(\sqrt{2E_b / N_0}), \quad (2.6)$$

где
$$F(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \exp(-t^2 / 2) dt$$
 .

Поскольку канал QPSK содержит два независимых канала BPSK, для QPSK формула (1.5) также применима. Зависимости $p(E_b/N_0)$ для некоторых видов модуляции приведены на рис. 2.11. Видно, что сигналы QAM и при одинаковой кратности с PSK обладают лучшей помехоустойчивостью. Это объясняется более равномерным расположением сигнальных точек по всему сигнальному пространству, что обеспечивает большие значения

минимального расстояния. С другой стороны, полная энергия QAM сигнала меняется во времени, поэтому он менее пригоден для работы в режиме насыщения спутникового ретранслятора.

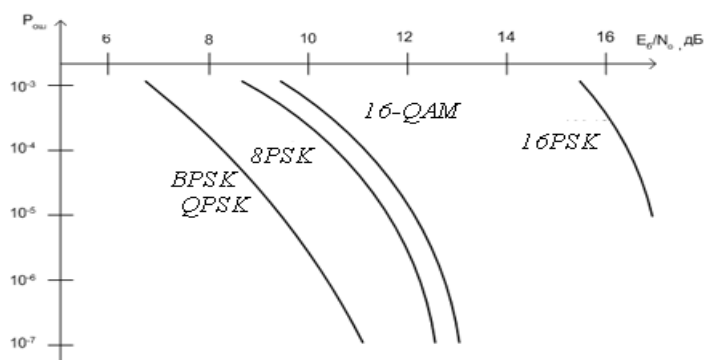


Рисунок 2.11 – Зависимость $P_{om}(E_b/N_b)$ для некоторых видов модуляции

Удельная **скорость передачи** на единицу полосы пропорциональна $\log_2 M$, где M - кратность модуляции. В частности, при одинаковых энергетических характеристиках скорость передачи QPSK на единицу полосы вдвое выше, чем у BPSK, что и обусловило ее широкое применение в системах связи, особенно спутниковых. PSK-8 обеспечивает скорость передачи втрое выше, чем BPSK, но требует более высокого отношения E_b/N_0 на входе демодулятора. Связь помехоустойчивости и информационной емкости при различных видах многопозиционной модуляции показана на рис.2.12, а зависимость скорости передачи от отношения сигнал/шум (SNR) при различных видах многопозиционной модуляции на рис.2.13.

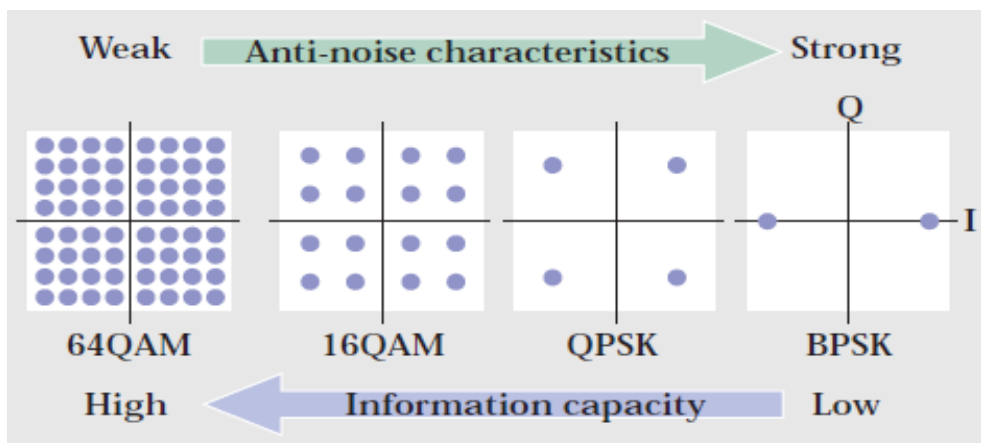


Рисунок 2.12 – Связь помехоустойчивости и информационной емкости при различных видах многопозиционной модуляции

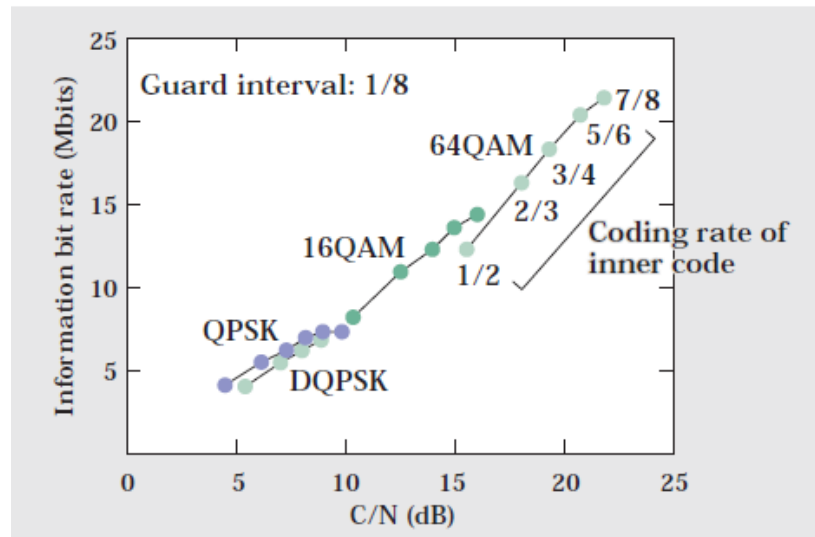


Рисунок 2.13 –Зависимость скорости передачи от отношения сигнал/шум (SNR) при различных видах многопозиционной модуляции, скоростях внутреннего канального кодирования защитном интервале $1/8T_s$

2.9 OFDM

До сих пор речь шла о фазовой и амплитудной модуляции одной несущей, используемых в цифровых спутниковых или кабельных системах телевизионного вещания, где основным источником ошибок являются тепловые шумы и уровень других мешающих сигналов относительно невелик. При переходе к эфирному цифровому вещанию выяснилось, что значительное влияние в метровом и дециметровом диапазонах волн оказывают отраженные сигналы, помехи от соседних по частоте цифровых и аналоговых сигналов, селективные замирания.

Основным разрушающим фактором для цифрового канала становятся помехи (интерференция) от многолучевого приема, при котором в декодер поступают две (или несколько) одинаковые по характеру чередования символов, но сдвинутые по времени последовательности. Если задержка одного из лучей становится равной или больше половины длительности символа, происходит резкий рост цифровых ошибок, вплоть до полного разрушения канала.

Специально для борьбы с помехами при многолучевом приеме было предложено использовать COFDM модуляцию (Coded Orthogonal Frequency Division Multiplexing - ортогональное частотное разделение каналов с кодированием), варианты которой использовались еще в пятидесятые годы прошлого века в тропосферной связи для борьбы с замираниями, связанными с многолучевостью распространения радиосигналов.

COFDM является цифровым методом модуляции, при которой используется большое количество близко расположенных ортогональных поднесущих. Каждая поднесущая модулируется по обычной схеме модуляции (например, [квадратурная амплитудная модуляция](#)) на низкой символьной скорости, сохраняя общую скорость передачи данных, как и у обычных схем модуляции одной несущей в той же полосе пропускания. На практике сигналы COFDM формируются путем использования быстрого обратного преобразования Фурье (FFT^{-1} на рис.2.14), а демодулируются с применением прямого преобразования Фурье (FFT на рис.2.15).

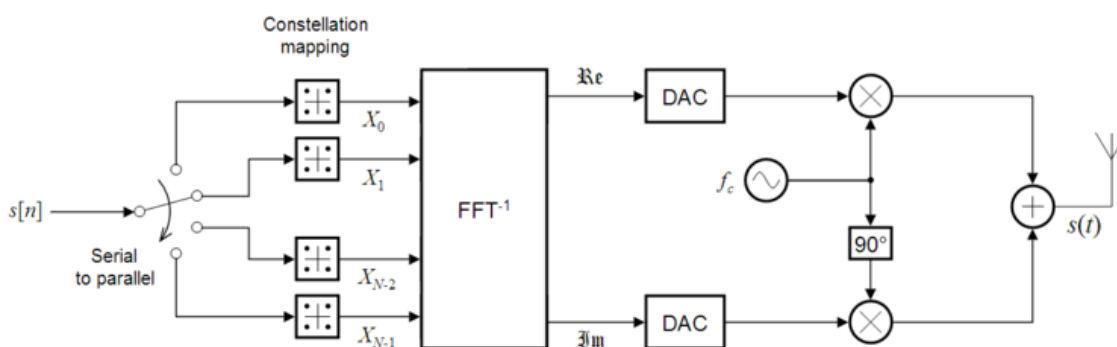


Рисунок 2.14 – Формирование сигналов COFDM с использованием быстрого обратного преобразования Фурье

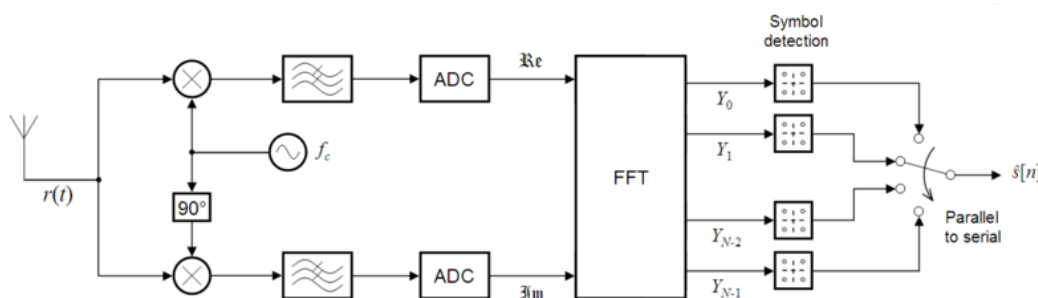


Рисунок 2.15 – Демодулируются сигналов COFDM с применением прямого преобразования Фурье

Основным преимуществом COFDM по сравнению со схемой с одной несущей является её способность противостоять сложным условиям в канале. Например, бороться с затуханием в области ВЧ в длинных медных проводниках, узкополосными помехами и частотно-избирательным затуханием, вызванным многолучевым характером распространения, без использования сложных фильтров-эквалайзеров. Канальная эквализация упрощается вследствие того, что COFDM сигнал может рассматриваться как множество медленно модулируемых узкополосных сигналов, а не как один быстро модулируемый широкополосный сигнал. Низкая символьная скорость делает возможным использование защитного интервала между символами, что позволяет справляться с временным рассеянием и устранять межсимвольные искажения (МСИ).

Основная идея OFDM - разделение последовательного цифрового потока на большое число низкоскоростных потоков, передаваемых на отдельных ортогональных несущих. Набор несущих, переносящих компоненты цифрового потока, называется символом OFDM. Благодаря большому числу несущих длительность символа в каждом из параллельных потоков оказывается в тысячи раз больше, чем в исходном последовательном потоке. Такая большая длина символа обеспечивает хорошую защиту от межсимвольных искажений, обусловленных интерференцией, так как отраженные сигналы чаще поражают не весь, а лишь часть символа. Степень защиты может быть значительно увеличена, если ввести защитный интервал D - промежуток времени, в течение которого оценка значения символа в декодере не производится. Обычно защитные интервалы размещаются в конце символов и не превышают $1/4$ длительности активной части символа.

Благодаря введению защитных интервалов декодер игнорирует часть разрушающих эхо-сигналов.

Уникальным свойством COFDM с защитными интервалами, во многом определившим выбор этого вида модуляции в Европе, следует признать возможность организации синхронной сети вещания, когда большое число ТВ передатчиков работают строго на одной частоте и синхронно передают одинаковые символы COFDM. Если расстояние между передатчиками выбрано таким образом, что разница в запаздывании сигналов от соседних передатчиков не превышает длительности защитного интервала, приемники в сети не испытывают помех от наложения сигналов.

Ключевой вопрос в построении системы COFDM - выбор основных параметров: числа индивидуальных несущих на символ, их частотного разнеса, значений защитного интервала, метода модуляции несущих, метода синхронизации. Все параметры взаимосвязаны и выбираются путем компромисса.

Частотный разнос между несущими зависит от способа выделения в демодуляторе отдельных несущих. Если применить традиционное разделение с помощью полосовых фильтров, частотный разнос между модулированными несущими пришлось бы выбирать таким, чтобы их соседние боковые полосы взаимно не перекрывались. Это условие можно выполнить, выбрав величину частотного разнеса равной $\Delta f \geq 2/T_c$, однако при этом эффективность использования радиоспектра будет невысокой. Для COFDM выбран более эффективный метод на основе ортогональных преобразований. Напомним, что две модулированные несущие называются ортогональными, если интеграл от их произведения на периоде длительности символа равен нулю. Для выполнения условий ортогональности нужно, чтобы частотный разнос между несущими был постоянен и равен $1/T_c$. Тогда на центральной частоте каждой несущей спектральные компоненты всех остальных несущих проходят через 0 (рис.2.16) и не мешают демодуляции на приеме. Взаимные помехи от соседних несущих будут равны нулю, несмотря

на то, что их соседние боковые полосы взаимно перекрываются. Это позволяет очень эффективно, близко к теоретическому пределу, использовать полосу частот ТВ канала и вдвое повысить удельную скорость передачи по сравнению с фильтровым методом. Выбранный метод разделения объясняет появление определения "ортогональный" в названии схемы модуляции.

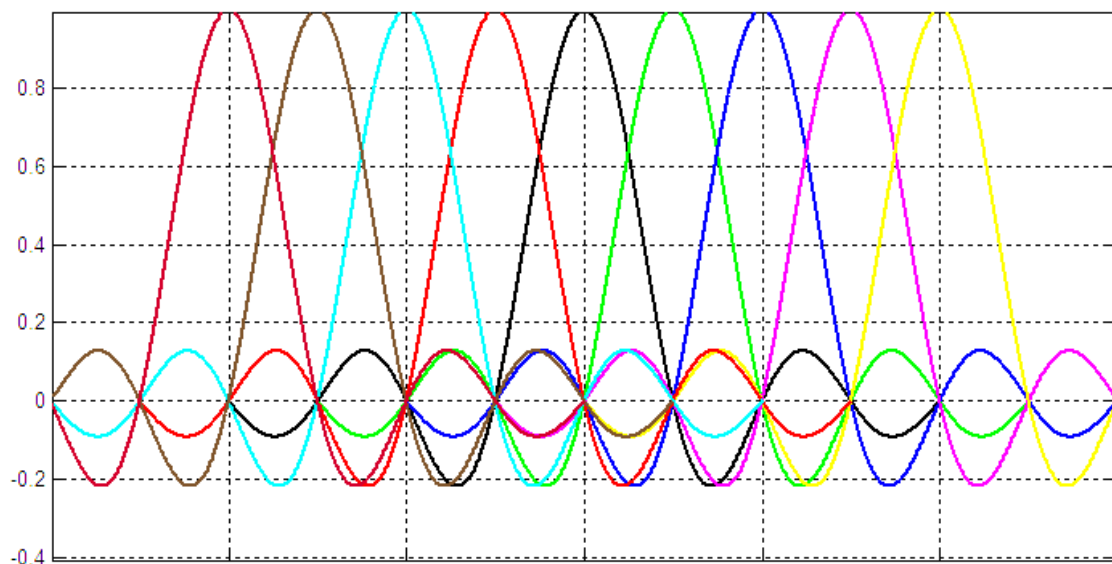


Рисунок 2.16 – Ортогональность отдельных несущих в спектре COFDM

Не менее важно для существования метода и определение "Coded" (кодированный). Если ограничиться использованием некодированной OFDM, из-за селективных замираний на некоторых частотах соответствующие несущие не будут декодироваться, давая всплески шума и ошибки по битам. Некодированная OFDM так же чувствительна к помехам в совпадающем и соседних каналах, как одночастотная модуляция. Например, узкополосная помеха, поражающая несколько несущих из 1000 в OFDM, приводит к вероятности ошибки 10^{-3} на уровне мощности помехи на 20 дБ ниже того, который вызвал бы аналогичную частоту ошибок в одночастотной системе. Применение помехоустойчивого кодирования повышает устойчивость передачи и делает прием более надежным. COFDM система, защищенная сверточным кодом с $R = 1/2$ и $K = 7$, при поражении 64 несущих из 512 показала ухудшение характеристик всего на 3 дБ.

Вопросы для самоконтроля

1. Чем отличается цифровое телевидение от аналогового?
2. Каковы основные принципы совместимости монохромного и цветного телевидения?
3. Как вычисляется частота дискретизации для ЦТС?
4. Как рассчитывается скорость передачи двоичных символов для ЦТС?
5. Как связана полоса спектра частот цифрового сигнала со скоростью передачи двоичных символов?
6. Какие способы уменьшения полосы частот канала могут быть применены в цифровом телевидении?
7. Какие цифровые устройства применяются в цифровом телевидении?
8. Какими путями достигается высокая помехоустойчивость цифрового телевидения?
9. Какие методы сжатия применяются для ЦТС?
10. В чём заключаются основные отличия стандарта MPEG-2 и MPEG-4.
11. В каком из стандартов предусмотрено применение фрактального сжатия.

Литература к разделу 2

1. Обработка изображений и цифровая фильтрация. Под ред. Т. Хуанга: Пер. с англ. — М.: Мир, 1979
2. Herve Benoit. Digital television : satellite, cable, terrestrial, iptv, mobile tv in the dvb framework.— 3rd ed. Focal Press is an imprint of Elsevier 30 Corporate Drive, Suite 400, Burlington, MA 01803, USA Linacre House, Jordan Hill, Oxford OX2 8DP, UK.
3. Ватолин Д., Ратушняк А., Смирнов М. Юкин В. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео// -М. – Диалог-МИФИ. -2003. - 384 с.

4. Bonomi, M. «Multimedia and CD ROM: An Overview of MPEG and JPEG» CD ROM Professional, November 1991, P. 38-40
5. Про сжатие видео — Введение [Электр. ресурс]. – Режим доступа: <http://habrahabr.ru/post/111244/>
6. Ричардсон Я. Видеокодирование H.264 и MPEG-4 – стандарты нового поколения.// - М. – Техносфера. – 2005. - 368 с.
7. Фитисов В. Использование кодека H.264 в вещании [Электр. ресурс]. – Режим доступа: <http://www.videoprotest.com/stati/obmen-opytom-2/item/219-использование-кодек-h264-в-вещании>
8. Тропченко А.А., Молчанов В.А. Особенности сжатия цветных изображений JPEG-подобными алгоритмами// Научно-технический вестник СПбГУ ИТМО/ - Вып. 32. - СПб. - СПбГУ ИТМО. – 2006. - с.22-26.