Кодирование речи методом RPE/LPC -LTP

1. RPE – LTP –кодер на 16 кбит/с

В 1990 г. предполагалось ввести в эксплуатацию Европейскую цифровую подвижную систему радиосвязи, в которой будет использоваться кодирование речевого сигнала со скоростью 16 кбит/с.

Разработка кодера производилась в 7 Европейских странах, а также в США и Канаде.

Были разработаны следующие системы:

* адаптивное дифференцирование ИКМ – ADPCM;
* адаптивное кодирование преобразованием – APC;
* кодирование с линейным предсказанием с возбуждением от остатка RELP–LPC;
* кодирование с линейным предсказанием с возбуждением от регулярных импульсов – RPE-LPC;
* кодирование с линейным предсказанием с многоимпульсным возбуждением – MPE-LPC,
* субполюсное кодирование –SBC –APCM.

 В конце разработки были проведены сравнительные испытания всех кодеров. Испытания проводились на 7 языках. В результате испытаний были отобраны два кодера:

* RPE (Regular–Pulse Excitation) - линейное предсказание с возбуждением от регулярных импульсов с долговременным предиктором LTP (Long Term Predictor)
* MPE–LTP -линейное предсказание с многоимпульсным возбуждением с долговременным предиктором LTP.

RPE– алгоритм предполагает, что сигнал остатка в линейном предсказании представляется последовательностью прореженных регулярных импульсов, но с большим числом импульсов в кадре, чем в многоимпульсном возбуждении MPE.

RPE кодеры менее сложные, однако качество речи при их использовании недостаточно хорошее из-за наличия в сигнале тонального шума, который получается в речевом сигнале в процессе высокочастотной регенерации.

В противоположность RPE – кодеру, кодер с многоимпульсным возбуждением MPE создает отличное качество речи, но является достаточно сложным.

Компромиссом между этими двумя вариантами является RPE–LTP кодер, т. е. линейное предсказание с возбуждением от регулярных импульсов и с долговременным предиктором - LTP.

В передающей части кодера производится кратковременный LPC анализ, долговременный LTP анализ и кодирование регулярных импульсов RPE – кодером (рисунок 1).

Коэффициенты отражения кратковременного предсказания получают по методу Берга для РФ 8-го порядка.

В кратковременном LPC анализе производится выделение коэффициентов отражения , преобразование их в коэффициенты логарифма площади (log-area-ratios), кодирование и передача их на прием.

Коэффициенты квантуют следующим образом: при i равном 1 и 2; 3 и 4; 5 и 6; 7 и 8 число бит на коэффициент соответственно равно 6; 5; 4; 2.

Итого, на 8 коэффициентов отводится 36 бит в кадре длительностью 20 мс.

Рисунок 1. Структурная схема кодера на 13 кбит/с.

В приемнике коэффициенты вновь преобразуются в коэффициенты отражения , которые затем используются для формирования инверсного решетчатого фильтра.

На выходе кратковременного LPC –анализатора появляется сигнал остатка, который поступает на долговременный LTP –анализатор.

Долговременный предиктор LTP размещается после кратковременного. Делается это для устранения периодичности, которая еще сохраняется в сигнале остатка кратковременного предиктора.

Такое размещение предикторов является наиболее приемлемым с точки зрения получения лучшего качества речи. Долговременный предиктор характеризуется выражением

 (1)

Коэффициенты отражения долговременного предсказания определяются также методом Берга для РФ 3-го порядка. На передачу каждого коэффициента отводится 3 бита в кадре. Коэффициенты предсказания предиктора определяются путем минимизации энергии остатка предсказания.

Взвешивающий фильтр с передаточной функцией

используется для корректировки формантных областей в спектре остатка предсказания относительно уровня шума квантования. Осуществляется это путем выбора γ.

Оптимальное значение γ определено путем прослушивания. Оно оказалось равным 0,7 … 0,9.

При этом воспринимаемое значение шума квантования становится минимальным.

Длительность импульсной характеристики составляет 11 выборок, при частоте дискретизации 8 кГц. Значения импульсной характеристики для соответствующих выборок с индексом представлены в таблице 1.


### Таблица 1 Значения импульсной характеристики

|  |  |  |  |
| --- | --- | --- | --- |
|  | 6 | 5(=7) | 4(=8) |
|  | 1,000000 | 0,700790 | 0,250793 |
|  | 2(=9) | 2(=10) | 1(=11) |
|  | 0,000000 | -0,045649 | -0,016356 |

Выход взвешивающего фильтра для каждого субкадра, длительностью 5 мс является , где номер выборки сигнала в субкадре с частотой дискретизации 8 кГц.

В соответствии с RPE алгоритмом, для уменьшения количества передаваемых дискретных отсчетов процесса, он подвергается предварительной обработке.

Дискретизированные с частотой 8 кГц отсчеты речи разбиваются на кадры, длительностью 20 мс, и 4 субкадра по 5 мс.

Субкадры процесса на выходе НЧ фильтра, длительностью 5 мс и состоящие из 39 отсчетов, подвергается децимации (прореживанию) в соотношении 1:3.

В результате получаются три выборки по 13 импульсов в каждой. Фазы этих последовательностей сдвинуты друг относительно друга на одну выборку (0,125 мс) (рисунок 2).

Далее производится выбор номера одной из этих трех последовательностей, обладающей с максимальной энергией, т. е.

В выбранной последовательности определяется импульс с максимальной амплитудой (масштабный) импульс .

В каждом 5 мс субкадре на передачу номера последовательности с максимальной энергией затрачивается 2 бита, а на передачу - 6 бит. кодируется по логарифмическому закону.

Кроме того, передаются амплитуды всех 13 импульсов выбранной последовательности с максимальной энергией.

При этом на передачу каждого импульса затрачивается 3 бита. На всю последовательность затрачивается бит в субкадре или бит в кадре.

Ниже приводится распределение битов по параметрам в 20 мс кадре: 8 коэффициентов ; 4 коэффициента ; 4 коэффициента ; 4 коэффициента ; 4 значения ; 4 значения всех 13 импульсов . Итого 260 бит/кадр.

Рисунок 2. Пример децимации и селекции импульсов

При частоте кадров 50 Гц общая информационная скорость составляет кбит/с. Для синхронизации и защиты от ошибок в канале связи отводится 3 кбит/с.

Кодер RPE-LTP-LPC обеспечивает высокое качество речи, которое незначительно снижается при 5% ошибок в канале связи и при отношениях сигнал/помеха 26 и 18 дБ.

Кодер может быть реализован на одном цифровом процессоре типа TMS320C25 с внешней памятью.

2. Структура декодера речи в стандарте GSM

Структурная схема декодера речи в стандарте GSM представлена на рисунке 3.

Рисунок 3. Структурная схема декодера речи стандарта GSM

Рассмотрим кратко структуру и работу декодера – синтезатора речи показанного на рисунке 3.

Из канала связи данные с помощью демультиплексора распределяются по различным блокам декодера. На RPE декодер поступают номер последовательности , максимальное значение импульса выборки, представляющей собой прореженный остаток предсказания.

Здесь отсчеты выборки масштабируются и дополняются нулями. Восстановленная таким образом выборка подается на LTP – синтезатор.

Его функции выполняет генератор на РФ третьего порядка с передаточной функцией .

На него подаются с демультиплексора коэффициенты отражения долговременного предсказания и период основного тона .

Синтезированный сигнал подается на LPC синтезатор, представляющий собой генератор кратковременного предсказания на РФ восьмого порядка с передаточной функцией .

Коэффициенты отражения на этот РФ поступают с демультиплексора через преобразователь коэффициента логарифма площади в по формуле

 (2)

Сигнал с выхода LPC–синтезатора для уменьшения шумов квантования поступает на постфильтр, на выходе которого получают декодированный речевой сигнал .

Кодеры с линейным предсказанием создают речь хорошего и отличного качества при скоростях передачи 9,6 кбит/с и выше. При скоростях ниже 9,6 кбит/с качество речи становится хуже из–за увеличения шумов квантования.

Для уменьшения их влияния осуществляется так называемая постфильтрация, с помощью которой изменяется спектр речевого сигнала так, что субъективно уменьшает восприятие шума квантования.

Постфильтр получается с помощью LPC – анализа, в котором содержится инверсный фильтр

 (3)

Рассмотрим взвешенный инверсный фильтр

 (4)

Коэффициент взвешивания не изменяет положение формантных частот, а изменяет только ширину формантных областей.

Взвешенный инверсный фильтр определяет полюса фильтра. Нули постфильтра определяет взвешенный инверсный фильтр вида

 (5)

При этих обозначениях передаточная характеристика постфильтра примет вид

 (6)

где и - коэффициенты взвешивания; и - порядок взвешивающих фильтров.

Эти параметры постфильтра обеспечивают необходимый вид спектральной характеристики постфильтра и формирование формантных областей.

При одних значениях области формант обостряются, при других – расширяются.

При значениях постфильтр имеет провалы в местах расположения формант, т. е. происходит искажение формантной структуры. Поэтому должно соблюдаться условие .

Постфильтр распределяет шумы квантования таким образом, что их величина становится больше в формантных областях и меньше между формантными областями в спектральных впадинах. Таким путем уменьшается субъективное восприятие шума.

В местах расположения формант шумы квантования маскируются речевым сигналом.

Но одновременно постфильтр искажает речевой сигнал. Параметры постфильтра выбираются так, чтобы не допустить больших искажений речи и по возможности уменьшить шумы квантования.

Параметры постфильтра и были определены экспериментально прослушиванием речи на выходе кодера.

Они оказались равными =0.95, =0.5…0.7. При этих значениях и получено повышение сегментального отношения сигнал/шум на 7…8 дБ и повышение разборчивости речи.

Таким образом, постфильтрация позволяет не только улучшить качество звучания, но и повысить разборчивость речевого сигнала на выходе кодера.