

УДК 004.934

ИСПОЛЬЗОВАНИЕ ЛИНЕЙНОГО ПРЕДСКАЗАТЕЛЯ РЕЧИ В ПРОГРАММНОЙ МОДЕЛИ НИЗКОСКОРОСТНОГО ВОКОДЕРА ДЛЯ ПЕРЕДАЧИ РЕЧИ ПО ГИДРОАКУСТИЧЕСКОМУ КАНАЛУ СВЯЗИ

Литвинцева А. В., Оболонин М. А.

ФБГОУ ВПО «Волгоградский государственный технический университет», Волгоград, Россия (400005, Волгоград, пр. им. Ленина, 28), e-mail: Angelinka_lav@mail.ru

Данная статья посвящена выбору вида низкоскоростного вокодера для создания программной модели для передачи речи по гидроакустическому каналу связи, отличающегося большим качеством передаваемой речи и большей естественностью ее звучания. Были рассмотрены различные виды вокодеров, выделены их достоинства и недостатки. Также рассмотрено использование линейного предсказателя речи на приемной стороне вокодера и особенности распространения звуковой волны в гидроакустической среде. В результате проведенного исследования за основу модели был выбран гармонический вокодер, снабженный линейным предсказателем речи на приемной стороне. С его помощью возможна передача сигнала с низкой скоростью, сохраняя при этом его параметры, характеризующие речь диктора. Эта модель позволяет контролировать длительность передаваемых и предсказываемых отрезков речи, что позволит выявить параметры для получения наибольшей эффективности использования линейного предсказателя в данной модели.

Ключевые слова: низкоскоростной вокодер, преобразование Фурье, линейный предсказатель, гидроакустический канал связи.

LINEAR PREDICTION OF SPEECH IN SOFTWARE MODEL OF LOW SPEED VOCODER OVER HYDROACOUSTIC COMMUNICATION CHANNEL

Litvintseva A. V., Obolonin M. A.

Volgograd State Technical University, Volgograd, Russia (400005, Volgograd, Lenin Avenue, 28), e-mail: Angelinka_lav@mail.ru

This article is about choosing the kind of low-speed vocoder to create a software model of voice transmission over hydro acoustic channel, which has better quality of speech and more natural sounds. We have considered different types of vocoder, highlighting their advantages and disadvantages. Also linear prediction of speech was examined for using in receiver part of vocoder and characteristics of the sound wave in sonar environment. As result of our work harmonic vocoder with linear prediction of speech in receiver was chosen for base of software model of low speed vocoder of voice transmission over hydro acoustic channel. It is possible to transmit signal with reduced rate while maintaining speech characteristics of a speaker. This model takes advantage in control of transmitted and predicted voice intervals, which can be useful to determine vocoder parameters to improve effectiveness of linear prediction of speech.

Key words: vocoder voice, speech transmission, hydroacoustic communication channel, Fourier transform.

Введение

Вокодер – устройство синтеза речи на основе произвольного сигнала с богатым спектром. В настоящее время существуют как аппаратные, так и виртуальные реализации данного устройства. Данное устройство позволяет обеспечить большую пропускную способность канала связи, производя сжатие потока информации, а также позволяет закодировать передаваемую информацию, что особенно важно в настоящее время, в связи с активным использованием и продолжающимся развитием техники связи, такой как беспроводная мобильная связь.

1. Виды вокодеров

Виртуальные вокодеры активно используются фирмами – производителями музыкального оборудования. Вокодер как музыкальный эффект позволяет перенести свойства одного (модулирующего) сигнала на другой сигнал, который называют носителем. В качестве сигнала-модулятора используется голос человека, а в качестве носителя – сигнал, формируемый музыкальным синтезатором или другим музыкальным инструментом.

Виртуальные вокодеры появились впоследствии на основе аппаратных, поэтому они часто имеют аналогичную структуру. Они состоят из передающей и принимающей части. Приемная сторона включает в себя синтезатор речи, который состоит из генератора тонального сигнала, генератора шумов и набора фильтров. Генератор тонального сигнала служит для воспроизведения гласных звуков, тогда как генератор шума – согласных [6].

Существуют следующие виды вокодеров:

1. Полосной вокодер.
2. Формантный вокодер.
3. Гармонический вокодер.
4. Корреляционный вокодер.
5. Гомоморфный вокодер.
6. Вокодер на основе линейного предсказателя.

Разрабатывается программная модель вокодера, то есть виртуальный вокодер, в основе которого лежит один из видов вокодеров.

Основными требованиями к вокодеру являются:

- возможность изменять частоту передаваемого сигнала так, чтобы сигнал проходил по гидроакустическому каналу связи с наименьшими потерями;
- отсутствие зависимости от языка речи;
- возможность контроля скорости передачи;
- несложная интерпретация структуры в программный вид.

Рассмотрим вокодеры некоторых видов и оценим их соответствие предъявленным требованиям.

Спектр речевого сигнала полосного вокодера на передающем конце разделяется узкополосными фильтрами (ПФ) на частотные полосы (спектральные каналы), в каждой из которых путем детектирования и сглаживания фильтрами низких частот определяется средняя интенсивность сигнала. Величины средних интенсивностей (временные огибающие) передаются в аналоговой или импульсной форме по каналу связи. Кроме того, на передающем конце определяется высота основного тона с помощью выделителя основного тона (ВОТ) и характер спектра сигнала возбуждения (тональный или шумовой) с помощью выделителя тон-шум (ВТШ). Эти сигналы также передаются по каналу. На приеме сигнал,

характеризующий основной тон, управляет частотой генератора импульсов (ГИ), а переключение входных фильтров на выход ГИ или генератора шума (ГШ) осуществляется с помощью сигнала тон-шум. Широкополосный сигнал, созданный одним из генераторов (сигнал возбуждения), разделяется на частотные полосы гребенкой (комплексом) ПФ. С выхода последних частотные составляющие сигнала возбуждения подаются на модуляторы (М), в которых с помощью временных огибающих спектральных каналов осуществляется управление их амплитудами. Для устранения нежелательных продуктов модуляции, на выходе модуляторов включается еще одна гребенка полосовых фильтров. Полученный синтезированный сигнал приближенно отображает исходный естественный сигнал [1].

Основным недостатком полосных вокодеров является техническая сложность и громоздкость реализации, обусловленная большим количеством используемых фильтров. Качество восстанавливаемой речи может снижаться из-за того, что в полосе пропускания фильтра на тональных звуках может оказаться несколько гармоник основного тона и число таких гармоник может меняться во времени.

Принципы построения формантного вокодера во многом аналогичны принципам естественного речеобразования и приема речи. В формантном вокодере происходит выделение из речевого сигнала управляющих сигналов (сигнал – параметров), которые на приеме воздействуют на резонансные контуры и воспроизводят требуемую огибающую спектра. Для передачи формантных параметров необходимы полосы, более узкие, чем для полосных вокодеров, в отношении занимаемой ширины канала преимущество будут иметь формантные вокодеры, по сравнению с полосными [7].

Главным недостатком формантного вокодера является трудность точной и надежной идентификации формант. Также для осуществления возможности понижения скорости передачи необходимо уменьшать словарь. Немаловажен и факт, что форматный вокодер обладает повышенной чувствительностью к ошибкам канала связи. Потенциальный недостаток данных вокодеров заключается в том, что синтезированная на приемной стороне речь не будет содержать индивидуальных характеристик голоса говорящего.

Коэффициент компрессии гармонического вокодера в 1,7 раза выше, чем полосного. Спектральную огибающую речевого сигнала можно представить в виде суммы ортогональных функций. В этом случае, в отличие от обычных – полосного и формантного вокодеров, спектральная огибающая на приеме воспроизводится не по отдельным ординатам, а в виде суммы тех же ортогональных функций. В настоящее время наиболее полно разработан метод представления спектральной огибающей в виде суммы гармонических функций.

Для гармонического вокодера достаточно иметь канал связи, обладающий пропускной способностью, не большей 2000 бит / с. Однако гармонические вокодеры не нашли практического применения из-за отсутствия существенных преимуществ по сравнению с полосными или форматными вокодерами. Но для программной реализации структура гармонического вокодера более удобна, и те незначительные преимущества при аппаратной реализации становятся более весомыми.

Одним из недостатков полосных и частично гармонических вокодеров является интерференция между составляющими спектра, а также возникновение временных сдвигов между составляющими спектра в полосных вокодерах из-за применения полосовых фильтров в синтезаторах. Одной из попыток избежать влияния эффектов фазовых сдвигов является применение корреляционных методов анализа речевого сигнала и соответствующего синтеза. Этот метод близок к методу гармонических вокодеров с анализом огибающей энергетического спектра речи.

Информацию о спектре можно выразить функцией корреляции. В корреляционных вокодерах, как и в гармонических, используется преобразование Фурье. Между функцией корреляции и энергетическим спектром сигнала имеется вполне определенная связь. Это дает возможность по функциям корреляции находить ординаты спектральной огибающей и наоборот.

В методе, используемом в корреляционных вокодерах, используется конечная задержка. Измеренная функция корреляции оказывается обрезанной и в синтезированном колебании обычно имеются разрывы непрерывности.

Гомоморфная обработка сигналов представляет собой один из нелинейных методов обработки, который может эффективно применяться к сложным сигналам, например, к речевым. С учетом используемой в вокодерах модели системы голосообразования, речевой сигнал можно представить как временную свертку импульсной переходной характеристики голосового тракта с сигналом возбуждения. В частотной области это соответствует произведению частотной характеристики голосового тракта и спектра сигнала возбуждения. Наконец, если взять логарифм от этого произведения, то получим сумму логарифмов спектра сигнала возбуждения и частотной характеристики голосового тракта. Поскольку человеческое ухо практически не чувствительно к фазе сигнала, можно оперировать с амплитудными спектрами. В гомоморфных вокодерах используется обратное преобразование Фурье.

2. Особенности распространения звука в воде

Звуковые волны являются единственным видом излучения, способным распространяться в толще океана на многие тысячи километров. Световые лучи практически

полностью рассеиваются в морской воде на протяжении нескольких десятков метров, а радиоволны затухают на расстояниях порядка сотни метров. Звук в воде распространяется гораздо дальше, чем в воздухе. Потери при распространении можно рассматривать как сумму потерь на расширение фронта волны и потерь вследствие затухания [5].

Особый вид распространения наблюдается в том случае, когда импульсный сигнал затягивается во времени. При многолучевом распространении происходит увеличение длительности импульса по мере удаления от источника. Это явление играет особенно важную роль при распространении сигнала на большие расстояния в подводном звуковом канале [4].

Одним из наиболее важных факторов, влияющих на процесс распространения звуковых волн в океане, является неоднородность воды на разных глубинах, приводящая к зависимости скорости распространения от глубины погружения. С увеличением глубины поглощение уменьшается. Поглощение звука при распространении в воде зависит от его частоты. Поглощение звуковых волн в морской воде повышается с увеличением их частоты. Главная причина этого – специфические процессы диссоциации и рекомбинации ионов, растворенных в воде солей. Вследствие этого поглощение звука в соленой воде существенно больше, чем в пресной. На низких частотах оно невелико, однако с увеличением частоты растет довольно быстро. Так, энергия звуковой волны уменьшается в 10 раз при частоте в 50 гц на расстоянии 25 000 км, при частоте 500 гц – на расстоянии 800 км, при частоте 5000 гц – на расстоянии 25 км [3].

Существует такое явление, как подводный звуковой канал (ПЗК). Часть энергии, излучаемой источником в ПЗК, остается в канале и не претерпевает акустических потерь, связанных с отражением от поверхности и дна.

3. Программная модель вокодера

Таким образом, из вышеизложенного можно сделать вывод, что для основы для программного моделирования низкоскоростного вокодера для передачи речи по гидроакустическому каналу связи наиболее всего подходит гармонический вокодер, а также структура программной модели может быть дополнена линейным предсказателем речи. Гармонический вокодер имеет показатель компрессии выше, чем у полосного вокодера, не зависит от языка и не требует дополнительного словаря, как формантный вокодер. Он имеет структуру, хорошо поддающуюся моделированию, в отличие от корреляционного и гомоморфного вокодеров. Также немаловажен факт, что гармонический вокодер работает на низких скоростях, без ущерба для качества речи, в отличие от формантного вокодера, который хоть и способен работать на более низких скоростях, но только за счет уменьшения размера словаря, что сказывается на качестве речи.

Сложный не периодический речевой сигнал можно разбить на короткие интервалы, в которых этот сигнал можно считать периодическим. В случае с речью наиболее удобно брать интервал равный средней длительности форманты. Подвергнув полученный отрезок сигнала обработке с помощью преобразования Фурье, получаются коэффициенты, которые будут определять гармоники, входящие в сигнал. Сам сигнал можно представить в виде гармоник вида:

$$F(k) = A \cos(2\pi vt + \varphi),$$

где:

v – это частота для гармоники с индексом k .

По полученным коэффициентам преобразования Фурье определяются наиболее значимые гармоники, по сумме которых можно сделать предсказание поведения продолжения сигнала для первого вокализованного его отрезка на следующие отсчеты времени (рисунок 1), количество которых меньше, чем количество отсчетов в исходном сигнале.

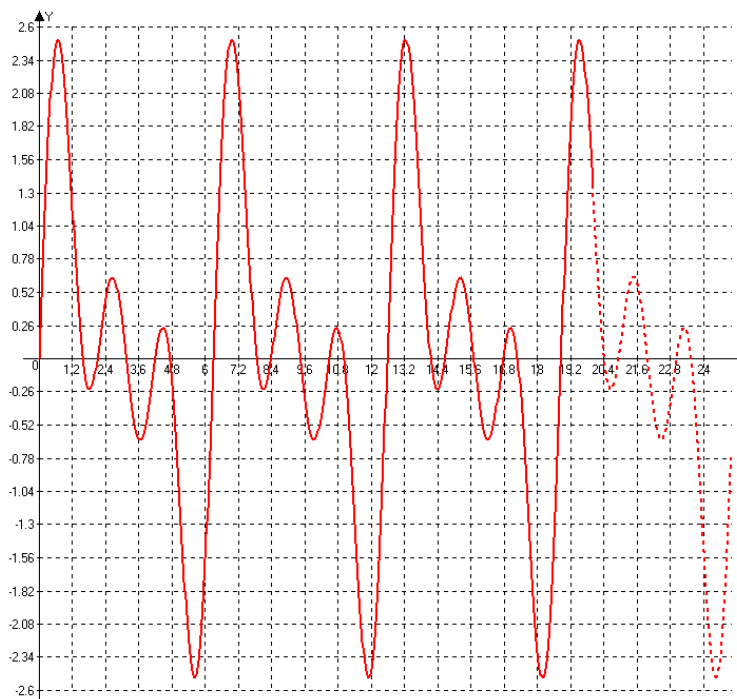


Рисунок 1. Сигнал с предсказанным продолжением

Для последующих отрезков производится аппроксимация функции на предыдущий отрезок и на пришедший.

Для невокализованных звуков производится определение принадлежности части сигнала к периодическим на заданном отрезке. В случае его периодичности он проходит ту же обработку, что и вокализованные сигналы, иначе – передается без изменений.

Таким образом, получается сигнал следующего вида, где пунктирной линией показаны предсказанные участки (рисунок 2). Как видно из рисунка, треть сигнала, полученного на выходе приемной стороны вокодера, является предсказанной.

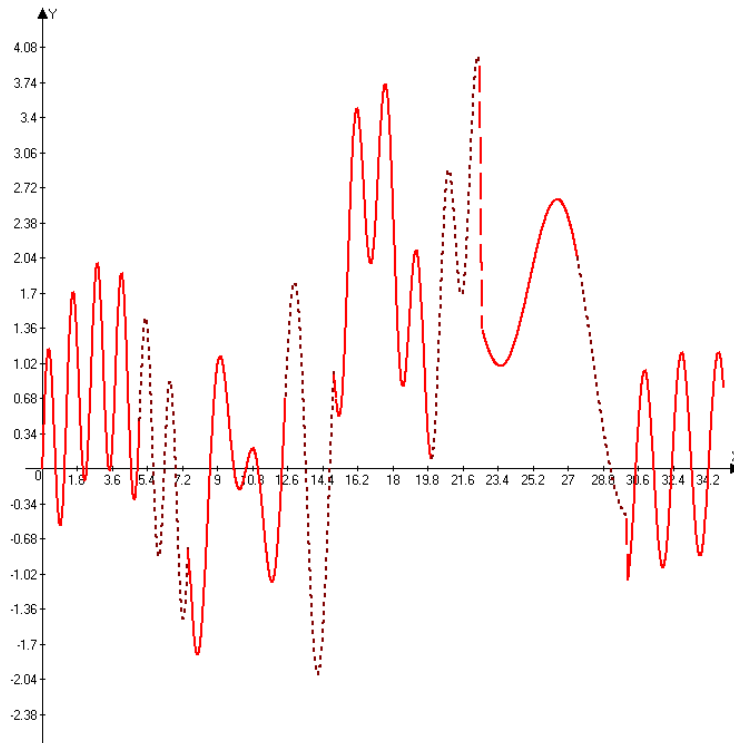


Рисунок 2. Полученный сигнал.

Данная обработка речи позволяет уменьшить скорость передачи, сохраняя при этом параметры сигнала, характеризующие речь диктора, что позволит повысить радиус его распространения в гидроакустической среде. В данном случае предсказание речи будет производиться аппроксимацией сигнала, но в случае, если встроить в линейный предсказатель обучаемую систему, то можно еще более повысить качество сигнала. Средняя длительность форманты, используемая в данной модели как длительность передаваемого отрезка, не имеет абсолютной величины. Эта величина равняется приблизительно 30 – 50 мс. [2]. При этом гласные звуки имеют большую длительность, чем согласные. В связи с этим данная модель снабжается настройкой длительности передаваемого интервала, а также возможностью передавать интервалы различной длительности для гласных и согласных звуков. Это позволит провести исследование и анализ результатов для выявления метода, который позволяет достичь лучших характеристик переданной речи.

Заключение

Таким образом, было принято решение создать программную модель вокодера на основе гармонического вокодера, принимающая часть которого снабжена линейным предсказателем. Сигнал подвергается кодированию и передается в цифровом виде, что позволяет контролировать частоту передачи, а также делает сигнал более

помехоустойчивым. Данная модель также позволит варьировать длительность передаваемых и предсказываемых отрезков речи, что позволит провести эксперименты для выявления эффективности применения линейного предсказания при различном соотношении исходной и предсказываемой части.

Список литературы

1. Акбулатов А. Ш. Вокодерная телефония. Методы и проблемы / Акбулатов А. Ш., Баронин С. П., Куля В. И. – М.: Связь, 1974.
2. Бондарко Л. В. Звуковой строй современного русского языка: Учеб. пособие для студентов пед. институтов по специальности «Рус. яз. и литература». – М.: Просвещение, 1977.
3. Корякин Ю. А. Корабельная гидроакустическая техника: Состояние и актуальные проблемы / Корякин Ю. А., Смирнов С. А., Яковлев Г. В. – СПб.: Наука, 2004.
4. Михальцев И. Е. Акустика и океанология // Вестник РАН. – 1960. – № 1.
5. Роберт Дж. Урик. Основы гидроакустики. – СПб.: Изд-во «Судостроение», 1978.
6. Сапожков М. А. Вокодерная связь / М. А. Сапожков, В. Г. Михайлов. – М.: Радио и связь, 1983.
7. Сапожков М. А. Речевой сигнал в кибернетике и связи. – М.: Гос. изд-во лит-ры по вопросам связи и радио, 1963.

Сведения о рецензентах:

Лукьянов Виктор Сергеевич, д-р техн. наук, профессор, зав. кафедрой «Электронно-вычислительные машины и системы» ВолгГТУ, г. Волгоград.

Камаев Валерий Анатольевич, д-р техн. наук, профессор, зав. кафедрой «Системы автоматизированного проектирования и поискового конструирования» ВолгГТУ, г. Волгоград.