

## АНАЛИЗ ВОЗМОЖНОСТИ ПЕРЕХОДА К СИСТЕМАМ С ПЕРЕМЕННОЙ СТРУКТУРОЙ ПРИ ЭФФЕКТИВНОМ КОДИРОВАНИИ РЕЧЕВОГО СИГНАЛА

Богачев Г.В., Афанасьев А.А., Илюшин М.В.

Академия ФСО России, e-mail: andreyska@mail.ru

Разработка эффективных способов и устройств сжатия и рационального представления видео- и звуковой информации является предпосылкой более эффективного использования каналов связи. Рост объемов информации, передаваемой по сетям электросвязи, сопровождается сохранением доминирующей роли речевых сообщений. Известно, что в системах связи объем телефонного трафика является основным, так как речь является наиболее естественной формой человеческого общения [1].

Разрабатываются новые высокоэффективные алгоритмы цифровой обработки речи различного назначения, возрастает количество приложений таких алгоритмов в самых разнообразных системах связи, появляются новые способы передачи речевой информации. Наиболее бурно развивающейся областью цифровой передачи речи являются системы мобильной радиосвязи (сотовой, транкинговой, спутниковой), системы Интернет-телефонии, в том числе с использованием радиоканалов [2].

В настоящее время в технике кодирования речевого сигнала используются в основном параметрически адаптивные алгоритмы кодирования речи на основе различных модификаций метода линейного предсказания, которые заключаются в различном представлении сигнала возбуждения [3, 4]. Основная задача, встающая перед разработчиками, – это качественное и компактное представление параметров характеризующих передаточную характеристику речевого тракта, а также сигнала возбуждения, аппроксимирующего ошибку линейного предсказания [5]. Основной особенностью адаптивных алгоритмов является обеспечение изменения параметров и/или структуры объекта управления (в данном случае - кодека) в соответствии с результатами анализа входных или выходных данных.

Анализ современных низкоскоростных кодеков речевых сигналов показывает, что в их развитии можно выделить несколько тенденций:

- повышение качества передачи на заданных скоростях;
- сохранение качества передачи при снижении скорости;
- расширение сфер применения кодеков речевых сигналов;
- широкое использование линейного предсказания и векторного квантования, как основных макропроцедур кодирования речи [2];
- усложнение алгоритмов обработки сигналов, за счет применения различных адаптивных алгоритмов [2,6].

Системы, основанные на процедуре линейного предсказания, относятся к классу адаптивных систем с переменными параметрами и показывают достаточно хорошие результаты в качестве кодирования. Однако данные алгоритмы имеют недостаток, связанный со значительной потерей качества синтезированной речи на участках с низкими корреляционными связями и нечувствительностью их ко многим статистическим и параметрическим характеристикам речевого сигнала [5]. Поэтому возникла потребность в улучшении качества кодирования речи за счет более глубокого анализа приходящего сигнала и использования новых подходов в области эффективного кодирования и адаптивных систем.

Фиксированное структурирование пространства кодируемых параметров и постоянство мощностей подпространств представлений (порядок предсказания для линейного предсказания), присущие существующим алгоритмам и выражающиеся в фиксированной структуре кодеков, не создают необходимых условий для максимального использования имеющейся априорной и апостериорной информации о речевом сигнале, что препятствует более глубокой оптимизации кодека.

В связи с вышеизложенным, перспективным направлением при создании новых систем низкоскоростного кодирования речи представляется разработка адаптивных алгоритмов кодирования речевых сигналов обеспечивающих функционирование кодека с переменной структурой [7]. Особенности построения таких систем были освещены в [8]. Основной задачей при работе такого кодека будет являться выбор рациональной структуры кодера и его параметров на этапе анализа и передача параметров выбранной структуры на декодер.

Исследуется подход к построению системы с переменной структурой, основанный на разделении сигнального пространства речевого сигнала на подпространства на основе анализа фундаментальных свойств и характеристик речи, важных для субъективного восприятия речи. Такой подход позволяет учесть не только статистические свойства речевого сигнала, но и особенности восприятия отдельных его фрагментов ухом человека. Классические методы структурного синтеза систем, описанные в (7), трансформированы применительно к задаче разработки адаптивных кодеков речевых сигналов на основе системы с переменной структурой с учетом вводимых ограничений.

Разработка адаптивных систем с переменной структурой для кодирования речи направлена на повышение качества передачи речевых сигналов, предлагается использовать не только адаптивное изменение параметров системы компрессии, но и адаптивное изменение ее структуры.

Естественно, что осуществление управления в системах с переменной структурой затруднено [6]. Если объект управления имеет структуру, скачкообразно изменяющуюся в случайные (квазислучайные) моменты времени, то это приводит к трудностям и особенностям оптимизации систем управления. Особенность состоит в том, что в процессе оптимизации необходимо оценивать состояние структуры объекта и менять алгоритм управления в зависимости от сложившейся ситуации. Основной задачей оптимизации является определение алгоритма управляющего устройства (оптимальных управлений) при имеющейся информации о состоянии объекта управления. Управление может осуществляться в условиях помех при передаче информации по каналам связи с ошибками. Здесь возможны два случая: в первом моменты изменения структуры могут быть известны точно, во втором за счет помех в канале эти моменты определяются с определенной вероятностью. В первом случае специфика задачи заключается в квазислучайном характере моментов переключения, а во втором – в необходимости оптимальных апостериорных оценок состояний структуры системы кодирования.

Применение полного перебора при дискретизации структуры системы кодирования или оптимизации структуры эволюционным методом при наложении ограничений на ее вариативность требует достаточно высокой вычислительной сложности от системы, осуществляющей физическую реализацию алгоритма кодирования. Данные подходы не оправдывают себя, так как в современных системах связи кодеки часто являются составной частью более сложного устройства, в связи с чем затруднено выделение достаточно большого дополнительного ресурса вычислительной системы для обработки в условиях реального масштаба времени.

Анализ основных стандартов кодирования речевых сигналов на основе метода линейного предсказания, выявил наиболее часто используемый класс кодеков данного типа – CELP (Code Excited Linear Prediction – линейное предсказание с возбуждением от кода).

Такие кодеки имеют недостаток, связанный со значительной потерей качества синтезированной речи на участках с низкими корреляционными связями и нечувствительностью их ко многим статистическим и параметрическим характеристикам речевого сигнала, в связи с чем, теряется потенциально возможное улучшение качества кодирования речевого сигнала. Подобные алгоритмы, применяемые в технике кодирования речи в настоящее время, относятся к классу адаптивных систем с изменяемыми параметрами.

Разработка способа кодирования речевых сигналов на основе системы с переменной структурой направлена на устранение данных недостатков, улучшение адаптивных свойств системы обработки.

Данное изменение структуры производится за счет использования апостериорных данных анализируемого сегмента на основе более глубокого изучения характеристик приходящего речевого сигнала. В модели кодека, основанной на таком подходе, более точно учитываются свойства речевого сигнала, за счет необходимой адаптации структуры кодирующего устройства на всех этапах обработки речевых сигналов.

Основными элементами системы с переменной структурой для кодирования речевых сигналов являются: выделители параметров классификации, устройство идентификации структуры (классификатор); устройство идентификации параметров; устройства, обеспечивающие реализацию адаптивных макропроцедур компрессии речи; устройства, предназначенные для дополнительной обработки речевого сигнала.

Для кодирования сегментов речи используются основные макропроцедуры низкоскоростного кодирования речевых сигналов [2]. В предлагаемом подходе структура кодека изменяется таким образом, что происходит переход от спектрально-временного параметрического кодирования на основе модели линейного предсказания к одной из разновидностей кодирования формы сигнала – векторному квантованию и наоборот.

В соответствии с вычисленными статистическими и параметрическими характеристиками обрабатываемого сегмента выбирается та структура кодирующего устройства, которая обеспечивает минимум искажений речевого сигнала на сегменте анализа.

При работе кодека с переменной структурой функционируют два адаптивных контура управления: первый производит оптимизацию структуры кодирующего устройства, второй – оптимизацию передаваемых параметров для выбранной структуры.

В такой системе поиск локально оптимальной структуры будет осуществляться на основе анализа апостериорных данных сегмента (пространство параметров классификации сегментов), а также априорных данных кодера и декодера о кодируемых сообщениях.

Таким образом, перспективным направлением развития адаптивных кодеков речевых сигналов является разработка систем с переменной структурой, то есть систем, в которых реализуются процедуры структурно-параметрической адаптации. Систему адаптивного структурно-параметрического кодирования речи целесообразно разрабатывать на основе следующих принципов:

- для усовершенствования выбирается класс CELP-кодексов речи, как наиболее широко используемый в системах связи и имеющий широкие возможности для развития;
- структурную перестройку кодека необходимо производить по входному сигналу путем деления обрабатываемых сегментов на классы, которые формируются с учетом особенностей восприятия речи

человеком; методы структурной перестройки системы, описанные в литературе, в кодеках речи применять нецелесообразно вследствие высокой их сложности и невозможности учета в них особых свойств восприятия речи человеком;

- параметрическую адаптивную обработку целесообразно реализовывать по принципу "анализа через синтез".

В системе с переменной структурой для кодирования речевых сигналов заложены принципы, которые позволяют улучшить качество кодирования речи, как по объективным, так и по субъективным показателям.

#### Литература

1. Вемян Г.В. Передача речи по сетям электросвязи. - М.: Радио и связь, 1985. – 272с.
2. О.И.Шелухин, Н.Ф.Лукьянцев Цифровая обработка и передача речи. М., Р и С, 2000г. – 456 с.
3. Коротаев Г. А. Анализ и синтез речевого сигнала методом линейного предсказания.// Зарубежная радиоэлектроника. – 1990. – № 3. – С. 31 – 51.
4. Коротаев Г. А. Некоторые аспекты линейного предсказания при анализе и синтезе речевого сигнала. // Зарубежная радиоэлектроника. – 1991. – № 7. – С. 3 – 31.
5. Маркел Дж. Д., А.Х. Грэй Линейное предсказание речи – М.: Связь, 1980. – 308с.
6. Уидроу Б., Стирнз С. Адаптивная обработка сигналов. – М.: Радио и связь.–1989, 450с.
7. Бухалев В.А. Распознавание, оценивание и управление в системах со случайной скачкообразной структурой. – М. Наука, 1996. – 288с.
8. Казаков И.Е. Статистическая динамика систем с переменной структурой.–М.:Наука,1977.–416с.

