Лабораторная работа №2

Кодирование линейных спектральных параметров

В предыдущей лабораторной работе решалась задача построения линейного фильтра для эффективного предсказания речевого сигнала. Представление фильтра его коэффициентами называют представлением во временной области. В данной работе будет рассмотрено другое описание фильтра, основанное на вычислении так называемых линейных спектральных параметров. Хотя эти параметры однозначно вычисляются по коэффициентам фильтра, это альтернативное представление играет большую роль в технике кодирования речевых сигналов. Основная причина состоит в том, что линейные спектральные параметры фильтра порядка p представляют собой упорядоченный набор из p чисел, принимающих значения из конечного интервала $[0, f/2]$, где f – частота отсчетов речевого сигнала. Экспериментально установлено, что диапазон изменения значений каждого из спектральных параметров относительно невелик, кроме того, на вокализованных участках речи спектральные параметры изменяются относительно медленно во времени по сравнению с параметрами предсказывающего фильтра. Благодаря перечисленным свойствам, квантование спектральных параметров оказывается предпочтительнее квантования коэффициентов фильтра.

Цель настоящей работы – изучение преобразования коэффициентов фильтра в набор линейных спектральных параметров, обратного преобразования, моделирование и исследование характеристик квантователя линейных спектральных параметров.

Представление фильтра с помощью линейных спектральных параметров

Рассмотрим спектральные характеристики авторегрессионной модели из лабораторной работы №1. Если $\{x(n)\} = x(1), x(2), ..., x(N)$ – последовательность из N отсчетов, то ее дискретное преобразование Фурье определяется по формуле

$$X(e^{j\omega}) = X(z)\Big|_{z=e^{j\omega}} = \sum_{n=1}^{N} x(n)e^{j\omega n}.$$

Пусть теперь $x(n)$ – авторегрессионный процесс, тогда преобразование Фурье входной и выходной последовательностей связаны соотношением

$$X(e^{j\omega}) = Y(e^{j\omega})H(e^{j\omega}),$$

где $H(e^{j\omega})$ – амплитудно-фазовая характеристика фильтра: $H(e^{j\omega}) = 1/A(e^{j\omega})$. Модуль амплитудно-фазовой характеристики $\left|H(e^{j\omega})\right| = 1/\left|A(e^{j\omega})\right|$ называют амплитудно-частотной характеристикой фильтра. Частоты, соответствующие нулям полинома $A(z)$, $z_i = r_i e^{j\omega_i}$, $i = 1,2..., p$, называют *формантными частотами*.

Как упоминалось в лабораторной работе 1, речевой сигнал можно моделировать как отклик линейной системы с переменными параметрами на соответствующий возбуждающий сигнал (для звонких звуков источнику возбуждения соответствует квазипериодическая последовательность импульсов, для шипящих звуков – шумовой сигнал). Преобразование Фурье речевого сигнала равно произведению преобразований Фурье возбуждающего сигнала и импульсного отклика голосового тракта. В частности, спектр, соответствующий периодическому возбуждающему сигналу, является линейчатым. Частотная характеристика голосового тракта является сравнительно гладкой функцией частоты; поскольку голосовой тракт представляет собой полость, то, в первую очередь, он характеризуется акустическими резонансами, соответствующими резонансным частотам этой полости, которые и называются формантными. Спектр речевого сигнала образуется перемножением линейчатого спектра возбуждающего сигнала и спектра, соответствующего голосовому тракту, и, следовательно, также является линейчатым, а его огибающая характеризует передаточную функцию голосового тракта.

При синтезе речи в рамках системы анализа-синтеза речевых сигналов линейный фильтр, описывающий голосовой тракт, может быть задан как во временной, так и в частотной области. В случае задания фильтра во временной области задача сжатия речевого сигнала сводится к квантованию и последующей передаче (хранению) коэффициентов фильтра $\{a_i\}, i = 1,2,..., p$. При переходе в частотную область квантованию с последующей передачей (хранением) подвергаются *линейные спектральные параметры*, или частотные параметры речевого сигнала, определенным образом связанные с его формантными частотами и получаемые по описанному ниже алгоритму.

Алгоритм вычисления линейных спектральных параметров

Поясним идею описанного ниже преобразования. Полином $A(z)$ степени p имеет в точности p корней. Если фильтр устойчив, его корни могут быть представлены в виде точек, лежащих внутри единичной окружности на плоскости комплексных чисел. Поскольку коэффициенты полинома – вещественные числа, корни его – попарно комплексно сопряженные. На первом шаге преобразования по полиному $A(z)$ строятся два вспомогательных полинома степени $p+1$ – суммарный и разностный, каждый из которых имеет корни, лежащие на единичной окружности. Каждый из полиномов имеет тривиальный вещественный корень, что позволяет понизить порядок полиномов до p. Далее, поскольку все корни – точки единичного круга комплексной плоскости, они могут быть заданы величинами фаз. Комплексно сопряженным корням соответствуют противоположные фазы. Это наблюдение позволяет простыми преобразованиями каждое из двух уравнений степени p свести к уравнению степени $p/2$ относительно косинусов фаз корней. Вычислив арккосинусы, получаем ровно p чисел φ_i, $i = 1,..., p$, лежащих в интервале $[0, \pi]$. Эти числа определяют частоты, соответствующие полюсам передаточной функции фильтра. Нижняя граница диапазона (число 0) соответствует постоянной составляющей сигнала, верхняя граница (число π) соответствует максимальной частоте $f/2$, где f – частота отсчетов речевого сигнала. Поэтому выполняется нормировка $\Phi_i = \varphi_i f/(2\pi)$, $i = 1,..., p$. Полученные значения называют линейными спектральными параметрами (ЛСП). Ниже детально, по шагам, описывается процесс вычисления ЛСП по коэффициентам фильтра, а затем восстановления коэффициентов фильтра по известным значениям ЛСП.

Шаг 1. Построение суммарного $P(z)$ и разностного $Q(z)$ полиномов на основе исходного полинома $A(z)$. Представим полином $A(z)$ в виде

$$A(z) = A(1)z^p + A(2)z^{p-1} + \ldots + A(p)z + A(p+1),$$

где $A(1) = 1, A(2) = -a_1, \ldots, A(p+1) = -a_p$. Построим суммарный полином

$$P(z) = P(1)z^{p+1} + P(2)z^p + \ldots + P(p+2)$$

порядка $p+1$ по следующему правилу:

$$P(1) = 1;$$

$$P(p+2) = 1;$$

$$P(k) = A(k) + A(p+3-k), \quad k = 2, \ldots, p/2+1;$$

$$P(k) = P(p+3-k), \; k = p/2+2, \ldots, p+1.$$

Например, для фильтра порядка $p = 2$ имеем

$$A(z) = A(1)z^2 + A(2)z + A(3),$$

тогда

$$P(z) = 1 + (A(3) + A(2))z + (A(3) + A(2))z^2 + z^3.$$

Построим разностный полином

$$Q(z) = Q(1)z^{p+1} + Q(2)z^p + \ldots + Q(p+2)$$

порядка $p+1$ по следующему правилу:

$$Q(1) = 1;$$

$$Q(p+2) = -1;$$

$$Q(k) = A(k) - A(p+3-k), \quad k = 2, \ldots, p/2+1;$$

$$Q(k) = -Q(p+3-k), \; k = p/2+2, \ldots, p+1.$$

Например, для фильтра порядка $p = 2$ имеем

$$Q(z) = -1 + (A(3) - A(2))z + (A(2) - A(3))z^2 + z^3.$$

Нетрудно видеть, что

$$A(z) = (P(z) + Q(z))/2z.$$

Шаг 2. Понижение порядка на 1. Строим суммарный полином $PL(z)$ порядка p из полинома $P(z)$ по правилу:

$$PL(1) = 1,$$

$$PL(k) = P(k) - PL(k-1), \quad k = 2, \ldots, p/2+1,$$

$$PL(k) = PL(p+2-k), \quad k = p/2+2, \ldots, p+1.$$

Отметим, что при выполнении условий устойчивости полученный полином имеет только комплексные корни и эти корни всегда лежат на единичной окружности.

Например, для фильтра порядка $p = 2$

$$PL(z) = 1 + (A(3) + A(2) - 1)z + z^2,$$

что эквивалентно понижению порядка полинома $P(z)$ путем деления на $(z+1)$.

Построим разностный полином $QL(z)$ порядка p по правилу:

$$QL(1) = 1;$$

$$QL(k) = Q(k) + QL(k-1)\,, \quad k = 2,..., p/2+1\,;$$
$$QL(k) = QL(p+2-k)\,, \quad k = p/2+2,..., p+1\,.$$

Можно показать, что в случае устойчивого фильтра корни этого полинома – комплексные и лежат на единичной окружности. Для фильтра порядка $p = 2$

$$QL(z) = 1 + (A(2) - A(3) + 1)z + z^2\,,$$

что эквивалентно понижению порядка полинома $Q(z)$ путем деления на ($z-1$).

Шаг 3. Понижение порядка до $p/2$. Учитывая, что корни полиномов $PL(z)$ и $QL(z)$ лежат на единичной окружности, т.е. имеют вид $z = e^{\pm j\omega_i}$, где $i = 1,..., p/2$, можно понизить порядок уравнения, которое необходимо решить для нахождения корней $QL(z) = 0$ и $PL(z) = 0$. Очередной этап нахождения линейных спектральных параметров состоит в построении полинома $P^*(z)$ и полинома $Q^*(z)$ порядка $p/2$ по следующим формулам:

$P^*(p/2+1) = 2^{-p/2}(PL(p/2+1) - 2PL(p/2-1) + 2PL(p/2-3) - ...);$

$P^*(p/2) = 2^{-(p/2-1)}(PL(p/2) - 3PL(p/2-2) + 5PL(p/2-4) - ...);$

$P^*(p/2-1) = 2^{-(p/2-2)}(PL(p/2-1) - 4PL(p/2-3) + 9PL(p/2-5) - ...);$

. . .

$P^*(2) = PL(2)/2;$

$P^*(1) = 1.$

Формальное правило, по которому записаны коэффициенты в этих формулах – следующее. В формуле для $P^*(p/2+1)$ в скобках за коэффициентом 1 следует последовательность из чередующихся коэффициентов -2 и +2. Во всех следующих формулах первый коэффициент равен 1. Знаки следующих коэффициентов чередуются. Коэффициент при i-м члене каждой формулы при $i > 1$ равен по абсолютной величине сумме абсолютных величин $(i-1)$-го коэффициента данной формулы и i-го коэффициента предыдущей формулы. Например в формуле для $P^*(p/2-1)$ имеем 4=3+1, 9=4+5, далее последует 16=7+9 и т.д.

Коэффициенты полинома $Q^*(z)$ могут быть получены аналогично из коэффициентов полинома $QL(z)$. В частности, для фильтра порядка $p = 2$ получаем

$$P^*(z) = (A(3) + A(2) - 1)/2 + z\ , \quad Q^*(z) = (A(2) - A(3) + 1)/2 + z\,.$$

Шаг 4. Решаем уравнения $Q^*(z) = 0$ и $P^*(z) = 0$. Эти уравнения имеют $p/2$ вещественных корней вида $z_i = \cos\omega_i$, где $i = 1,2,..., p/2$. Для фильтра порядка $p = 2$ получаем

$$zp = -(A(2) + A(3) - 1)/2\,, \quad zq = -(A(2) - A(3) + 1)/2\,.$$

Для решения уравнений произвольного порядка можно использовать любой метод численного решения алгебраических уравнений. На практике часто используют следующий метод. Поиск корней полинома производится в два этапа. Сначала на интервале -1, 1 ищут приближенные значения корней $\mathrm{root}(j), j = 1,2,..., p/2$. Для этого интервал делится на n равных подынтервалов, величина n зависит от порядка фильтра. Например, для $p = 10$ выбирают $n = 128$. Наличие корня в подынтервале обнаруживается по изменению знака полинома. На втором этапе производится уточнение значения корня. Уточненное значение корня $\text{exroot}(j), j = 1,2,..., p/2$ вычисляется по формуле

$$\text{exroot}(j) = \text{root}(j) + 2F(i-1)/(n(F(i-1) - F(i)),$$

где i – номер интервала, на котором располагается корень с номером j; $\text{root}(j) = -1 + 2(i-1)/n$; $F(i)$ - значение полинома в конце i -го подынтервала.

Шаг 5. Обозначим через $zp_i, i = 1,2,..., p/2$ корни полинома $P^*(z)$ и через $zq_i, i = 1,2,..., p/2$ корни полинома $Q^*(z)$. Находим линейные спектральные параметры по формулам

$$\omega p_i = \arccos(zp_i),\ i = 1,...p/2;$$
$$\omega q_i = \arccos(zq_i),\ i = 1,...p/2.$$

Далее выполняется сортировка значений по возрастанию и их нормировка умножением на частоту дискретизации и делением на 2π.

Полученные спектральные параметры могут быть проквантованы векторно или скалярно и закодированы вместе с другими параметрами речевого сигнала.

На приемной стороне решается задача восстановления коэффициентов фильтра по квантованным значениям спектральных параметров по следующему алгоритму.

Шаг 1. Квантованные значения разбиваются на подмножества, одно из которых (значения с четными номерами) соответствует суммарному полиному, другое – разностному. Значения умножаются на 2π и делятся на частоту дискретизации. Вычисляется косинус каждого значения. Тем самым найдены корни полиномов $P^*(z)$ и $Q^*(z)$.

Шаг 2. Пользуясь обобщением теоремы Виета, можно выразить все коэффициенты полиномов $P^*(z)$и $Q^*(z)$ через их корни. Для фильтра порядка $p = 2$ получаем $P^*(2) = -zp, Q^*(2) = -zq, P^*(1) = 1, Q^*(1) = 1$.

Шаг 3. Пользуясь уже приведенными формулами, связывающими коэффициенты полиномов $P^*(z), PL(z)$ и $Q^*(z), QL(z)$, рекуррентно восстанавливаем коэффициенты полиномов $PL(z), QL(z)$, полагая $PL(1) = QL(1) = 1$. Для фильтра порядка $p = 2$ получаем

$$PL(1) = 1, PL(2) = -2zp, PL(3) = 1;$$
$$QL(1) = 1, QL(2) = -2zq, QL(3) = 1.$$

Шаг 4. Восстанавливаем суммарный и разностный полиномы $P(z), Q(z)$ порядка $p+1$ по правилам

$$P(1) = PL(1);$$
$$P(j) = PL(j) + PL(j-1),\quad j = 2,.., p/2+1;$$
$$P(J) = PL(p+3-j),\quad j = p/2+2,..., p+2;$$
$$Q(1) = QL(1);$$
$$Q(j) = QL(j) - QL(j-1),\quad j = 2,.., p/2+1;$$
$$Q(J) = -QL(p+3-j),\quad j = p/2+2,..., p+2.$$

Для фильтра порядка $p = 2$ получаем

$$P(1) = 1;\ P(2) = -2zp + 1 = A(2) + A(3);\ P(3) = P(2);\ P(4) = 1;$$
$$Q(1) = 1;\ Q(2) = -2zq - 1 = A(2) - A(3);\ Q(3) = A(3) - A(2);\ Q(4) = 1.$$

Таким образом, восстановлены полиномы

$$P(z) = 1 + (A(3) + A(2))z + (A(3) + A(2))z^2 + z^3 \;;$$

$$Q(z) = -1 + (A(3) - A(2))z + (A(2) - A(3))z^2 + z^3 .$$

Шаг 5. Восстанавливаем полином $A(z)$

$$A(z) = (P(z) + Q(z)) / 2z .$$

Нетрудно убедиться в том, что при $p = 2$ эта формула приводит к правильному результату.

ЗАДАНИЯ

Исходные данные – файл, исследованный при выполнении лабораторной работы № 1 и порядок фильтра линейного предсказания.

1. Дополнить программу подсчета коэффициентов фильтра, разработанную при выполнении лабораторной работы № 1, программой расчета линейных спектральных параметров и восстановления коэффициентов фильтра по спектральным параметрам.
2. Определить экспериментально диапазон изменения каждого из линейных спектральных параметров.
3. Составить программу скалярного равномерного квантования линейных спектральных параметров.
4. Подавая на вход синтезирующего фильтра идеальное возбуждение, исследовать зависимость субъективного качества синтезированного сигнала от точности квантования каждого из спектральных параметров.
5. Построить графики зависимости отношения сигнал/шум от числа уровней квантования каждого спектрального параметра.
6. Выбрав число уровней квантования для всех параметров одинаковым, построить зависимость отношения сигнал/шум от числа уровней квантования.
7. Основываясь на оценках энтропии квантованных значений спектральных параметров, построить зависимость отношения сигнал/шум от затрат на передачу спектральных параметров. Определить битовую скорость, при которой квантование спектральных параметров не вносит искажений в синтезированный сигнал.
8. **Задание для самостоятельного исследования.** Используя подход Линде–Бузо–Грея, построить неравномерный квантователь спектральных параметров. Оценить эффективность неравномерного квантования спектральных параметров по сравнению с равномерным.