

Исследование речевых сигналов с низкоскоростным кодированием на наличие избыточности

Redundancy analysis of speech signals with low rate encoding

Ключевые слова: избыточность – redundancy; коэффициент избыточности – redundancy coefficient; взаимная информация – mutual information; низкоскоростное кодирование речи – low-rate voice coding.

В статье рассматривается избыточность в сигналах с низкоскоростным кодированием речи. Демонстрируется зависимость коэффициента избыточности от способа представления сигнала. Представлен подсчет взаимной информации, присутствующей в сигнале.

This article describes redundancy in signals with low rate encoding of speech. It illustrates the correlation between the redundancy factor and the signal presentation method. The calculation of the signal's mutual information is also provided.

Речевой сигнал по своей природе обладает избыточностью. Для иллюстрации данного факта может быть рассчитан коэффициент избыточности, который является количественной оценкой избыточности источника сообщения и вычисляется по формуле [1]:

$$\chi = \frac{H_{\max}(X) - H(X)}{H_{\max}(X)} = 1 - \frac{H(X)}{H_{\max}(X)} \quad (1)$$

где $H(X) = -\sum_{i=1}^m p(x_i) \log p(x_i)$ – энтропия источника с алфавитом из m сообщений; $H_{\max}(X) = \log m$ – максимальная энтропия источника с алфавитом из m сообщений.

Для исследования свойств наименее подверженных сжатию речевых сигналов может быть использован сигнал с выхода кодера ИКМ, описанного в рекомендации ITU-T G.711. Несмотря на

ПОЛЯКОВ / POLYAKOV S.

Сергей Владимирович

(pron_rzhew@mail.ru)

Академия ФСО России, адъюнкт.

г. Орел

то, что в области электросвязи за основу, то есть за сигнал, не подверженный процедурам устранения избыточности, принят именно этот сигнал, используется выражение «наименее подверженный сжатию». Это обуславливается тем, что алгоритм, описанный в указанной рекомендации, реализует переход от равномерного квантования к квантованию неравномерному [3].

Речевой сигнал с выхода кодера представляет собой набор отсчетов x_i по 8 бит. Следовательно, максимальная энтропия источника составит $H_{\max}(X) = \log 256 = 8$ бит.

Как известно, выражение для оценки вероятности имеет следующий вид:

$$p(x_i) = \lim_{n \rightarrow \infty} \frac{C_i}{n}$$

где C_i – количество испытаний с положительным исходом (в данном случае число 8-ми битных отсчетов, полученных с выхода кодера и имеющих одинаковое значение);

n – общее число испытаний (общее число 8-ми битных отсчетов, полученных с выхода кодера).

Результаты расчетов представлены на рисунке 1.

Из графика видно, что значения на выходе рассматриваемого кодера не равновероятны, следовательно, избыточность в сигнале присутствует. Для расчета коэффициента избыточности все данные необходимо подставить в формулу 1.

Целесообразно выяснить, присутствует ли избыточность в цифровом потоке, только в случае, если кодер рассматривается как источник сообщений с алфавитом $m = 2^8$ символов.

Результаты расчетов, представленные в таблице 1, подтверждают тот факт, что независимо от того, как рассматривается цифровой поток с выхода кодера G.711 (отсчетами по 1, 2, 3, 4, 5 ... 8 бит),

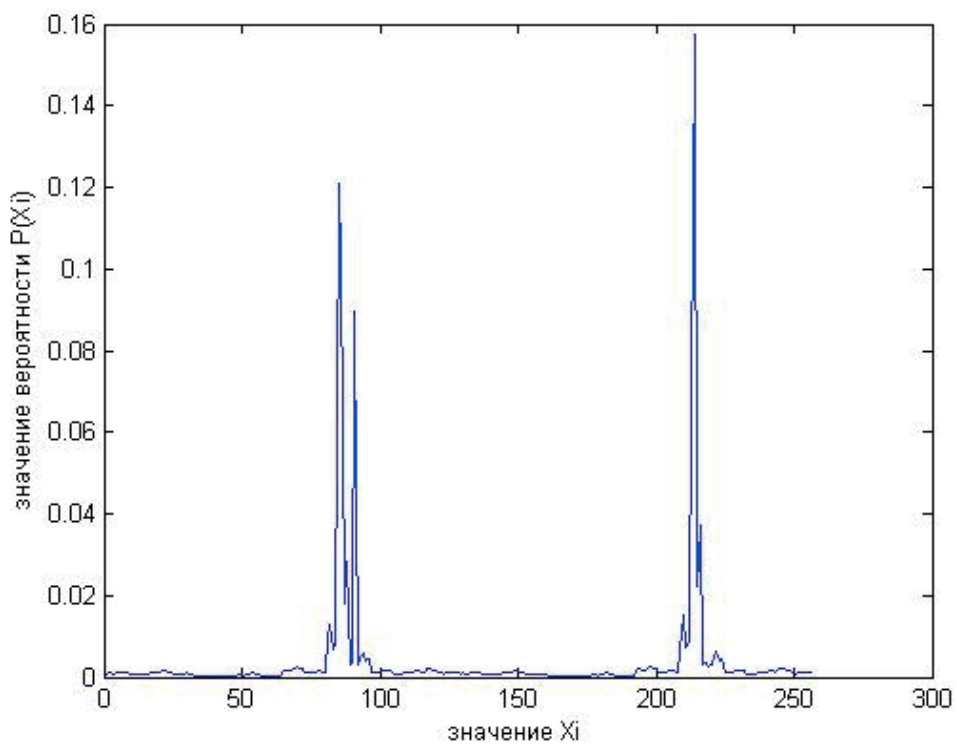


Рис. 1. Плотность вероятности появления значений x_i

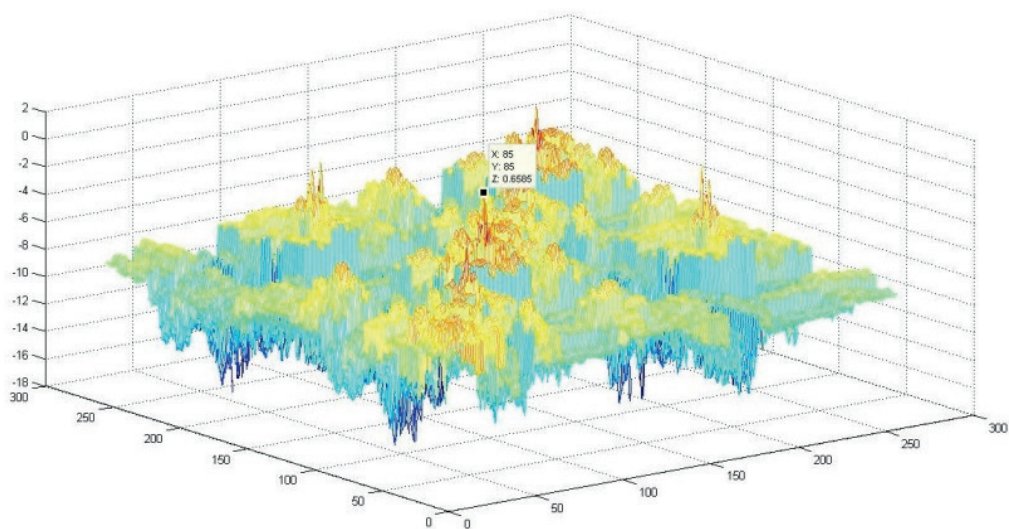


Рис. 2. Взаимная информация событий $X = x_k$ и $Y = x_{k+1}$ для размера алфавита источника 256 символов

Таблица 1

Величина коэффициента избыточности для различных кодеров

m	X										
	G.711		G.723		G.726			G.728	G.729		
	A-low	Mu-low	5,3 Кбит/с	6,3 Кбит/с	24 Кбит/с	32 Кбит/с	40 Кбит/с		6,4 Кбит/с	8 Кбит/с	11,8 Кбит/с
2 ¹	0,0015	0,0607	0,0029	0,0010	0,1888	0,0875	0,0649	0,0135	0,0019	0,0010	0,0018
2 ²	0,1801	0,0658	0,0032	0,0020	0,1972	0,1198	0,0877	0,0188	0,0031	0,0017	0,0031
2 ³	0,0803	0,0748	0,0032	0,0018	0,3128	0,1202	0,1002	0,0148	0,0033	0,0017	0,0035
2 ⁴	0,2333	0,1507	0,0054	0,0031	0,2170	0,2639	0,1094	0,0201	0,0055	0,0035	0,0041
2 ⁵	0,1055	0,0901	0,0049	0,0016	0,2238	0,1397	0,2297	0,0703	0,0038	0,0034	0,0044
2 ⁶	0,2018	0,1191	0,0049	0,0054	0,3257	0,1814	0,1213	0,0223	0,0050	0,0028	0,0052
2 ⁷	0,1245	0,1123	0,0047	0,0021	0,2342	0,1541	0,1273	0,0181	0,0045	0,0028	0,0056
2 ⁸	0,3324	0,3367	0,0098	0,0066	0,2398	0,2801	0,1325	0,0257	0,0177	0,0158	0,0072
	VoiceAge AMR-NB										
	4,75 Кбит/с	5,15 Кбит/с	5,90 Кбит/с	6,70 Кбит/с	7,40 Кбит/с	7,95 Кбит/с	10,2 Кбит/с	12,2 Кбит/с			
2 ¹	0,0165	0,0042	0,051	0,0099	0,0054	0,0069	0,0017	0,0006			
2 ²	0,0196	0,0059	0,0076	0,0121	0,0059	0,0083	0,0020	0,0009			
2 ³	0,0210	0,0057	0,0099	0,0127	0,0059	0,0095	0,0018	0,0017			
2 ⁴	0,0248	0,0109	0,0122	0,0155	0,0077	0,0119	0,0037	0,0029			
2 ⁵	0,0246	0,0071	0,0121	0,0125	0,0061	0,0093	0,0021	0,0025			
2 ⁶	0,0271	0,0109	0,0146	0,0199	0,0067	0,0165	0,0029	0,0032			
2 ⁷	0,0274	0,0207	0,0137	0,0137	0,0064	0,194	0,0025	0,0032			
2 ⁸	0,0318	0,0225	0,0301	0,0248	0,0109	0,0218	0,0059	0,0072			

можно установить факт наличия избыточности в данном цифровом потоке. Кроме этого, в таблице 1 приведены результаты расчета коэффициента избыточности для кодеров G.711 Mu-low, G.723, G.728, G.726, G.728, G.729, VoiceAge AMR-NB.

Одной из причин избыточности в исследуемом сигнале является наличие взаимной информации между отсчетами рассматриваемого сигнала. Другими словами, в событии, состоящем в том, что с выхода кодера будет получен отсчет x_k , содержится некоторое количество информации об отсчете x_{k+1} , который будет получен с выхода кодера на следующем шаге.

Выражение для оценки величины взаимной информации между событиями X и Y (информации о событии $X = a_i$, содержащейся в событии $Y = b_k$) имеет вид [2]:

$$I_{X;Y}(a_i; b_k) = \log \frac{P_{XY}(a_i | b_k)}{P_X(a_i)} \quad (2)$$

где $P_X(a_i)$ – априорная (безусловная) вероятность события $X = a_i$; $P_{XY}(a_i | b_k)$ – вероятность события $X = a_i$, при условии, что произошло событие $Y = b_k$.

Для данного случая выражение 1 примет следующий вид:

$$I_{X;Y}(x_{k+1}; x_k) = \log \frac{P_{XY}(x_{k+1} | x_k)}{P_X(x_{k+1})} \quad (3)$$

Необходимо оценить величины условных вероятностей для всех возможных значений x_k . Эти данные могут быть получены в результате натуральных экспериментов.

Результаты оценки представлены в виде трехмерной гистограммы на рисунке 2.

На рисунке 2 ось X – значения x_k ; ось Y – значения x_{k+1} ; ось Z – значения взаимной информации событий $X = x_k$ и $Y = x_{k+1}$.

Данные, представленные на гистограмме, свидетельствуют о наличии взаимной информации между значениями, получаемыми с выхода кодера. Кроме того, можно отметить, что величина взаимной информации не является постоянной для всех отсчетов.

Выводы

Цифровой поток с выхода речевого кодера обладает избыточностью. Независимо от того, какой способ представления цифрового потока с выхода кодера из вышеперечисленных выбран, наличие избыточности подтверждается значением коэффициента избыточности. Из расчетов, произведенных для различных типов кодеров, видно, что

избыточность в цифровом потоке зависит от применяемого кодера. Это связано с тем, что в основу работы кодеров положены различные алгоритмы кодирования. Наличие избыточности свидетельствует о наличии взаимной информации между цифровыми отсчетами, получаемыми с выхода кодера. Данный факт проиллюстрирован на примере подсчета взаимной информации между 8-ми битовыми отсчетами цифрового потока с выхода кодера G.711.

Литература

1. Теория электрической связи : учебное пособие / К.К. Васильев [и др.] ; под общей редакцией К.К. Васильева. – Ульяновск: УлГТУ, 2008. – 452 с.
2. Р. Галлагер. Теория информации и надежная связь. – М.: Советское радио, 1974. – 720 с.
3. ITU-T Recommendation G.711, Pulse code modulation (PCM) of voice frequencies, Geneva, 1988. – 10 с.