

ЭФФЕКТИВНОЕ ИСПОЛЬЗОВАНИЕ ИНТЕГРИРУЮЩЕЙ СПОСОБНОСТИ СЛУХА ПРИ ЦИФРОВОЙ ОБРАБОТКЕ СИГНАЛОВ

Стефанов А.М., Стефанова И.А.

Поволжская Государственная Академия Телекоммуникаций и Информатики
443090 г. Самара, Московское шоссе-77, кафедра Вычислительной Техники

Важнейшие характеристики любой системы цифровой обработки звуковых сигналов в значительной мере определяются эффективностью системы сжатия. В настоящее время повышение эффективности сжатия связано с обработкой звуковых сигналов в частотной области, что позволяет учитывать психофизиологические свойства органа слуха человека.

С точки зрения сжатия интерес представляет такое свойство слуха как интегрирующая способность – группирование частотных составляющих звука в определённые частотные полосы (группы) [1, 2]. Если расположить эти группы друг за другом без промежутков, то будет образован так называемый полосный ряд.

Вследствие разделения спектра звукового сигнала на частотные полосы слух реагирует не на общую мощность сигнала, а на мощность, сосредоточенную в отдельных частотных группах. При этом более интенсивные частотные группы при определённых условиях могут маскировать (подавление слухового восприятия) менее интенсивные группы. Применительно к сжатию это означает, что в группах меньшей интенсивности допустим больший уровень шума квантования, то есть допустимо уменьшение числа разрядов при кодировании соответствующих спектральных составляющих, что в свою очередь, снижает скорость передаваемого цифрового потока.

Экспериментально установлено, что ширина i – ой частотной группы ΔF_i зависит только от значения её средней частоты f_i [1]. Для практического использования этого свойства необходимо аналитическое описание зависимости ширины группы от значения средней частоты f . Это аналитическое описание должно соответствовать «тонкому» музыкальному слуху (восприятие звука в полосе 20-20000 Гц). Только тогда можно говорить о сохранении высокого качества воспроизведения при одновременном сжатии звукового сигнала.

Качественное представление о характере искомой зависимости даст кривая усреднённых значений ширины частотных групп, экспериментально измеренных при различных значениях средней частоты. При этом естественно сопоставить усреднённую кривую средней чувствительности слуха, характерной для большинства людей.

Известен полосный ряд [2], отображающий средние значения результатов экспериментальных измерений ширины частотных групп (среднее статистического «облака» данных). Для оценки степени приближения этого ряда к среднему авторами доклада был использован один из методов регрессионного анализа – линейная регрессия общего вида [3, 4]. При этом критерием точности приближения принята величина среднеквадратичного отклонения:

$$R^2 = \frac{\sum_{i=1}^n (\Delta F_{r_i} - \Delta F_s)^2}{\sum_{i=1}^n (\Delta F_i - \Delta F_s)^2}, \quad (1)$$

где ΔF_{r_i} – расчётное значение ширины группы в i -м эксперименте, n – общее число экспериментов, а ΔF_s – усреднённое по n значение ширины группы. В результате была получена функция с высокой степенью точности ($R^2 = 0,9993$) аппроксимирующая исходный полосный ряд. Однако, применение (2) непосредственно к статистическим данным дало $R^2 = 1,24$, что свидетельствует о недостаточной точности соответствия предложенного ряда среднестатистическим данным.

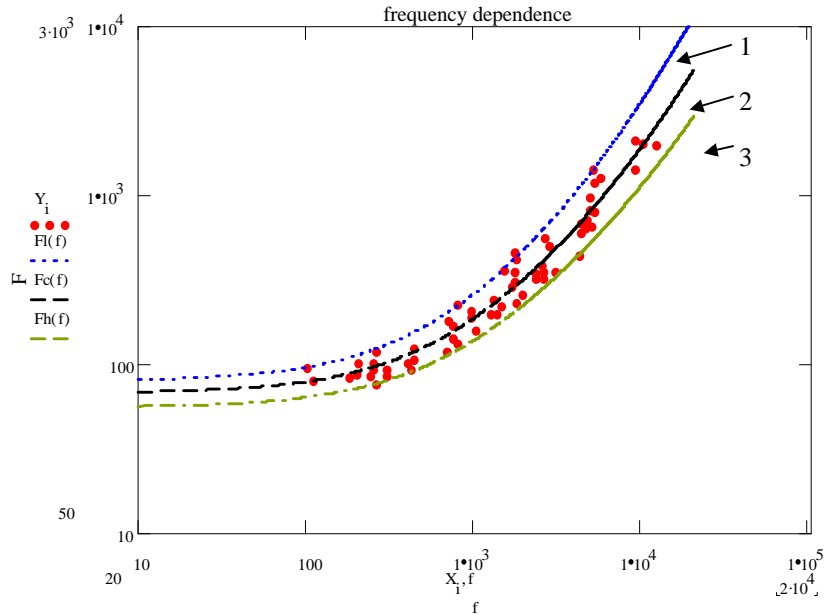


Рис.1. Зависимость ширины частотных групп F от центральной частоты f при разной чувствительности слуха

Используя тот же анализ, но уже на «облаке» статистических данных, удалось получить функцию, с максимальной степенью точности ($R^2 = 1$) аппроксимирующую усреднённые результаты измерений ширины группы при различных значениях частоты:

$$\Delta Fc(f) = 10^{-5,1} * f^2 + 0,1139 * f + 68 \quad (2)$$

$$\Delta Ft(f) = 0,933 * 10^{-5} * f^2 + 0,122 * f + 44,4 \quad (3)$$

Выражение (3) показывает, что основную роль в интегрирующем свойстве слуха играет квадратичная нелинейность, что хорошо согласуется с современным представлением о его нелинейных искажениях [2]. Рис.1 иллюстрирует семейство парабол, соответствующих различной чувствительности слуха. Точками показаны результаты измерений, а кривые 1, 2 и 3 соответствуют чувствительности уха ниже среднего Fl, средней чувствительности Fc и высокой чувствительности Fh, соответственно.

Таким образом, зависимость $\Delta Fh(f)$, соответствующую «тонкому» музыкальному слуху, следует искать в виде параболы, проходящей по нижнему краю статистического «облака» данных. Такая функция была получена авторами доклада в результате регрессионного анализа:

$$\Delta Fh(f) = 10^{-5,5} * f^2 + 0,0809 * f + 56 \quad (4)$$

Последнее выражение позволяет оценить число и ширину частотных групп для случая широкополосного сигнала. Результаты расчётов представлены в таблице 1.

Как видно из табл.1, в максимально слышимом диапазоне частот размещается 36 частотных групп, что прекрасно согласуется с экспериментальными результатами, полученными в [1] при вероятностном подходе к оценке максимально слышимого числа групп.

Таким образом, табл.1 обеспечивает возможность эффективного сжатия звуковых сигналов при использовании полосных методов их обработки [5], а выражение (4) – ортогональных базисов.

Таблица 1

№ группы	Центральная частота группы, Гц	Ширина группы, Гц	№ группы	Центральная частота группы, Гц	Ширина группы, Гц
1	50	60	19	2618	289
2	112.5	65	20	2922	319
3	180.5	71	21	3258	353
4	254.5	77	22	3630	391
5	334.5	83	23	4043	435
6	421.5	91	24	4502	484
7	516.5	99	25	5015	541
8	619.5	107	26	5589	607
9	731.5	117	27	6234	683
10	853.5	127	28	6961	772
11	986.5	139	29	7786	878
12	1132	152	30	8727	1003
13	1291	166	31	9805	1153
14	1465	181	32	11050	1336
15	1654	198	33	12500	1561
16	1862	218	34	14200	1842
17	2091	239	35	16220	2200
18	2342	263	36	18650	2665

Литература

1. Цвикер Э., Федкеллер Р. Ухо как приёмник информации. – М.: Связь, 1971.- 260 с.
2. Zwicker E. Psychoakustik. Springer – Veriag. – Berlin – Heidelberg – New York, 1982.
3. Ферстер Э., Ренц Б. Методы корреляционного и регрессионного анализа. М.: Финансы и статистика, 1983. – 210 с.
4. Носач В.В. Решение задач аппроксимации с помощью перспективных компонентов. – М.: БИНОМ, МИКАП, 1994. – 180 с.
5. Flanagan J.L. etal/ Speed Coding // IEEE Trans. – 1979. com-27. № 4. p.710-737.



EFFECTIVE USAGE OF THE INTEGRANT PHYSABILITY OF HEARING IN DIGITAL SIGNAL PROCESSING

Stefanov A.M., Stefanova I.A.

Povolzhskaya State Academy of Informatics and Telecommunications
443010 Samara, Moscovskoye Sh. – 77, Computing Department

The most important characteristics of any digital sound processing system are defined by the efficiency of data compression. Currently, the improvements of the compression efficiency are connected with sound signal processing in the frequency domain, which makes it possible to take into consideration psychophysiologic qualities of human ear.

Speaking about compression, the integrant physability of hearing, i.e. the ability to integrate frequency components of a sound into definite frequency bands [1,2], is of the utmost interest. If one disposes these bands in tandem without intervals then, the so-called, band set will be formed.

On account of disintegration of the sound signal spectrum into frequency bands our ear does not react to the whole power of a signal, but to the power concentrated in the definite frequency groups. Under certain circumstances, more intensive frequency groups can mask less intensive groups (the phenomenon of audio perception suppression). Thus, in the groups with lower intensity the higher level of noise quantization is allowed, i.e. it is possible to reduce the number of bits while coding the corresponding spectrum constituents, which, in its turn, reduces speed of the digital stream transmission.

It has been experimentally established that the width of “i” frequency group ΔF_i depends only on the value of its average frequency f_i [1]. In order to use this quality practically it is necessary to describe the dependence of the group width on the value of the average frequency f —analytically. The analytical description should correspond to a good ear (perception within the band of 20-20000 Hz). Only then you can talk of preserving a high-quality playback alongside sound signal compression.

The curve of the average values of the bandwidth, measured experimentally at different frequency values, reflects qualitatively the character of the looked over functional dependence.

We know the band set [2] reflecting average experimental values of the results of measuring the frequency groups width (average data “cloud”). In order to estimate the degree of correspondence of this set to an average set we used a method of regressive analysis – the linear regression of a general format [3,4]. The value of standard deviation is taken for the criterion of the degree of approximation:

where ΔF_i is the averaged value of the group width in “i” experiment, n is the total number of the experiments and ΔF_s is the averaged value of the group width.

$$R^2 = \frac{\sum_{i=1}^n (\Delta F_i - \Delta F_s)^2}{\sum_{i=1}^n (\Delta F_i - \Delta F_s)^2}, \tag{1}$$

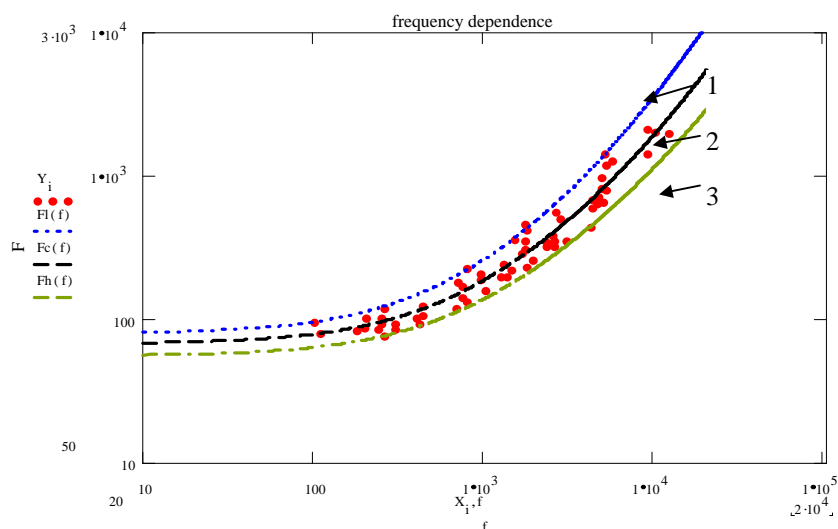


Fig1. The Dependence of the Frequency Group Width F on the Central Frequency f for Different Ear Sensitivities

As the result we got the following function:

$$\Delta F_c(f) = 10^{-5,1} * f^2 + 0,1139 * f + 68 \quad (2)$$

with the high degree of approximation ($R^2 = 0,9993$). However, its application (2) immediately to the statistic data showed $R^2 = 1,24$, which proves that the degree of correspondence of the suggested set to the average data is insufficient.

Using the same analysis, but taking into consideration the totality of the statistic data, we managed to get the function with the maximum degree of approximation ($R^2 = 1$) characterizing the average results of measurement of the group width at different frequency values.

$$\Delta F_t(f) = 0,933 * 10^{-5} * f^2 + 0,122 * f + 44,4 \quad (3)$$

The expression (3) shows that square non-linearity plays the most important role in the integrant physability to hear, which is well-related with the up-to-date ideas of non-linear distortions [2]. Fig.1 illustrates the family of parabolas corresponding to different degrees of ear sensitivity. The points show the results of measuring and the curves 1,2,3 correspond to the low ear sensitivity Fl, to the average ear sensitivity Fc and to the high sensitivity Fh accordingly.

So the dependence $\Delta F_h(f)$, corresponding to a good ear, is shown as the parabola running at the bottom of the totality of statistic data. We got this function as the result of regressive analysis:

$$\Delta F_h(f) = 10^{-5,5} * f^2 + 0,0809 * f + 56 \quad (4)$$

The last expression helps us to estimate the number and the width of frequency groups in case of a broadband signal.

The results of the research show that in the whole perceptible range there are 36 frequency groups, which agrees with the experimental results we have got (1) using the empirical means to evaluate the number of the maximum audible groups.

Thus, the results of the research provide the possibility of the effective sound signal compression with the help of band methods in signal processing (5) and formula (4) itself- with the help of orthogonal bases.

Literature.

1. Zwicker E., Fedkeller R. Ear as the Receiver of Information. – M.:Communications, 1971.-P. 260.
2. Zwicker E. Psychoakustik. Springer – Veriag. – Berlin – Heidelberg – New York, 1982.
3. Ferster E., Renz B. Methods of Correlative and Regressive Analysis. - M.:Finance and Statistics, 1983. – P.210.
4. Nosach V.V. The Solution of the Tasks of Approximation with the Help of Perspective Components. – M.: BINOM, MIKAP, 1994. – P.180.
5. Flanagan J.L. etal/ Speed Coding // IEEE Trans. – 1979. com-27. № 4. p.710-737.