

disadvantages of this type of analysis are investigated. A comparison is made of the results of revealing the reverse, gear and cyclic frequencies of the gearbox by the digital cepstral method with other digital methods for identifying the bases of harmonic series. Based on studies of the problems of digital cepstral processing of the energy spectra of signals, analysis of the advantages and disadvantages of this method of identifying the bases of scale, it is proposed, in order to increase the efficiency and effectiveness of cepstral analysis, that the energy spectra are not measured by the DFT method, but by the parametric DFT method (DFT-P). Briefly considered are the analytical properties of DFT-P. The existence of fast DFT-P (FFT-P) algorithms is proved. An experimental verification of the theoretical results obtained in the work was carried out by the method of numerical simulation.



МОДЕЛИРОВАНИЕ НЕЙРОСЕНСОРНОЙ ТУГОУХОСТИ НА ОСНОВЕ СУБПОЛОСНОЙ КОМПРЕССИИ ДИНАМИЧЕСКОГО ДИАПАЗОНА

асп. Порхун М.И., доц. Вашкевич М.И.

Белорусский государственный университет информатики и радиоэлектроники
ул. П. Бровки, 6, БГУИР, каф. ЭВС, 220013, Минск, Беларусь, e-mail: porhun@bsuir.by, vashkevich@bsuir.by

Описываются основные принципы построения компрессоров динамического диапазона для моделирования нейросенсорной тугоухости по аудиограмме. Рассматриваемая модель основана на совместном использовании субполосной декомпозиции звукового сигнала и применении компрессии в каждой полосе. Используя данный подход можно моделировать феномен усиленного нарастания громкости (рекруитмент), связанный с сужением слышимого динамического диапазона. Настройка параметров компрессии динамического диапазона выполняется на основании информации, извлекаемой из аудиограммы человека с нейросенсорной тугоухостью.

Введение. Одним из основных методов моделирования нейросенсорной тугоухости является субполосная декомпозиция звукового сигнала с последующим применением компрессоров динамического диапазона (КДД) в каждой полосе (рисунок 1) [1-3]. Разделение сигнала на субполосы осуществляется при помощи банка фильтров, как правило, согласованного с механизмом декомпозиции звука человеческим ухом. КДД отвечают за частотно-зависимое ослабление сигнала в полосах. Применение КДД позволяет отобразить узкий динамический диапазон остаточного слуха в широкий динамический диапазон нормального слуха, а также имитировать феномен усиленного нарастания громкости (рекруитмент), чего, как показывают исследования, вполне достаточно для моделирования нейросенсорной тугоухости [1,4]. Одной из ключевых задач подобных методов является корректная настройка КДД по аудиограмме человека с нейросенсорной тугоухостью.

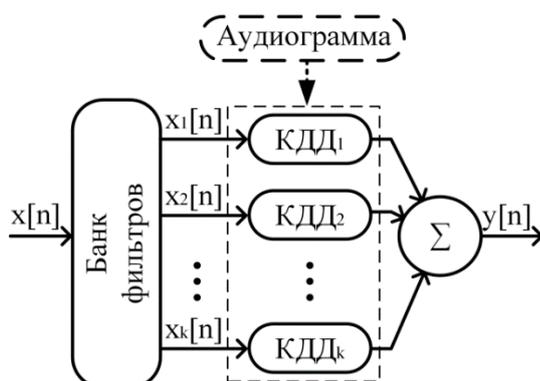


Рисунок 1 – Общая схема моделирования нейросенсорной потери слуха

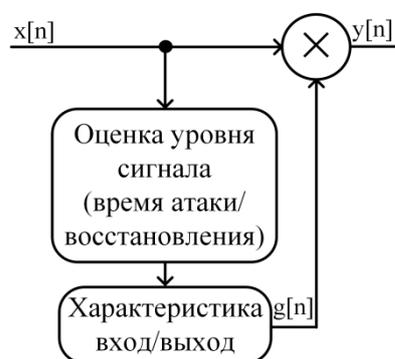


Рисунок 2 – Общая схема КДД

Компрессор динамического диапазона. Основная цель компрессии – автоматический выбор коэффициентов ослабления/усиления субполосных составляющих в зависимости от уровня мощности сигнала в текущей полосе. Компрессоры динамического диапазона (рисунок 2) характеризуются следующими параметрами: порог срабатывания (англ. *CT* – compression threshold), коэффициент сжатия

(англ. CR – compression ratio), время атаки (t_a) и время восстановления (t_r), а также коэффициент усиления/ослабления (G) [5]. Порог срабатывания определяет точку перегиба характеристики вход/выход КДД, тем самым определяя область, в которой алгоритм компрессии будет активен. Коэффициент сжатия определяет степень сжатия динамического диапазона. Время атаки задает промежуток времени между моментом установки входного уровня сигнала и временем, через которое выход оценщика уровня сигнала достигнет этого значения. Время восстановления же определяет промежуток времени между моментом спада уровня входного сигнала и временем, через которое данный уровень установится на выходе оценщика. Коэффициент усиления/ослабления показывает, как нужно изменить уровень входного сигнала в соответствии с характеристикой вход/выход [5]. Пример характеристики вход/выход КДД приведён на рисунке 3. Иллюстрация времени атаки и времени восстановления компрессора представлены на рисунке 4.

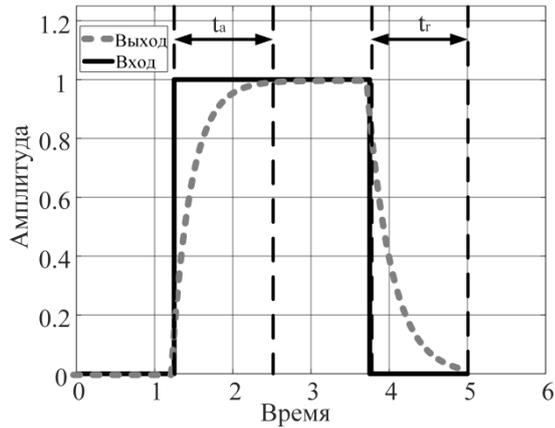
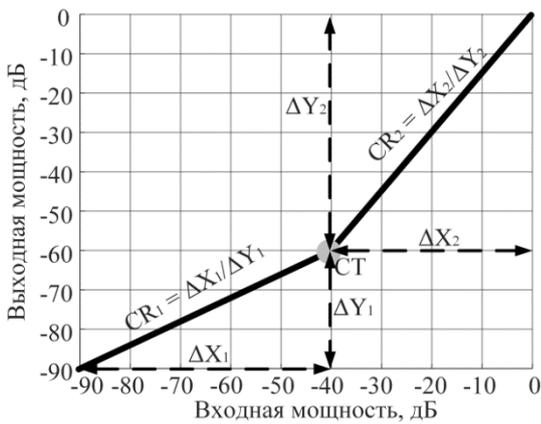


Рисунок 3 – Характеристика вход/выход КДД

Рисунок 4 – Время атаки и восстановления

Рекруитмент (феномен усиленного нарастания громкости). Известно [6], что здоровая улитка усиливает тихие звуки на 50...60 дБ, а громкие не усиливаются вовсе (рисунок 5). За величину коэффициента усиления отвечают наружные волосковые клетки. Появление рекруитмента (феномена усиленного нарастания громкости) связано с тем, что при увеличении потери слуха волосковые клетки теряют регулятивные свойства, что искажает процесс нормального усиления звуков: тихие звуки усиливаются крайне слабо по сравнению с громкими. Субъективно, для человека это проявляется в болевых или дискомфортных ощущениях, вызываемых громкими звуками.

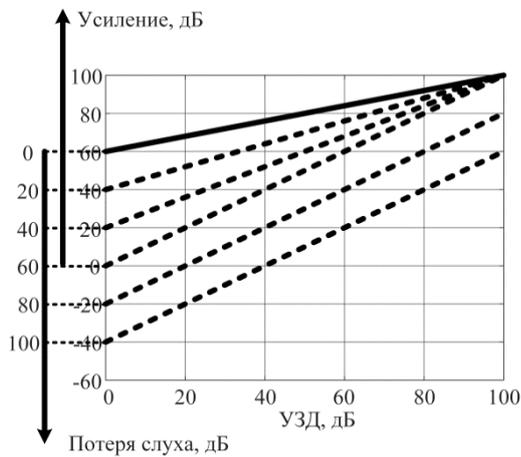


Рисунок 5 – Усиление в улитке

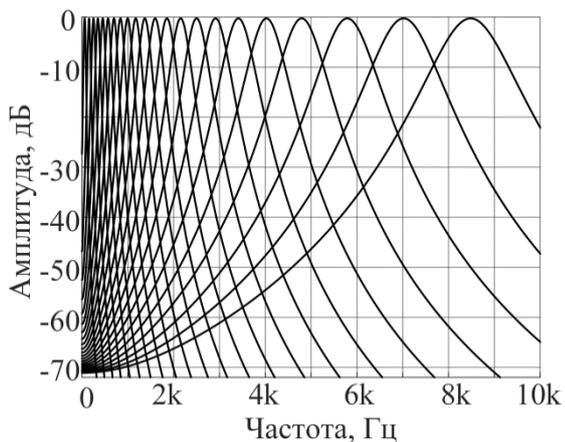


Рисунок 6 – АЧХ банка фильтров

Настройка КДД. Настройка КДД для конкретной субполосы выполняется по следующим параметрам: порог рекруитмента (T_R), порог слышимости в норме (T_N) и порог слышимости тугоухого человека (T_I).

Аудиограмма (рисунок 7) является основным источником информации при моделировании нарушений слуха. Поскольку для каждой частотной полосы строится свой компрессор, то для настройки КДД необходимо знать пороги слышимости на центральных частотах банка фильтров (рисунок 6) (f_c). Для их определения требуется интерполировать аудиограмму слабослышащего человека и абсолютный порог слышимости (в соответствии со стандартом ISO 226 [7]) на сетку центральных частот банка фильтров. Затем

необходимо рассчитать порог слышимости тугоухого человека (T_I) на центральных частотах банка фильтров (рисунок 8):

$$T_I(f_c) = T_N(f_c) + HL(f_c) \quad (1)$$

где $T_N(f_c)$ – порог слышимости для слуха в норме на частоте f_c , $HL(f_c)$ – уровень потери слуха из аудиограммы на частоте f_c .

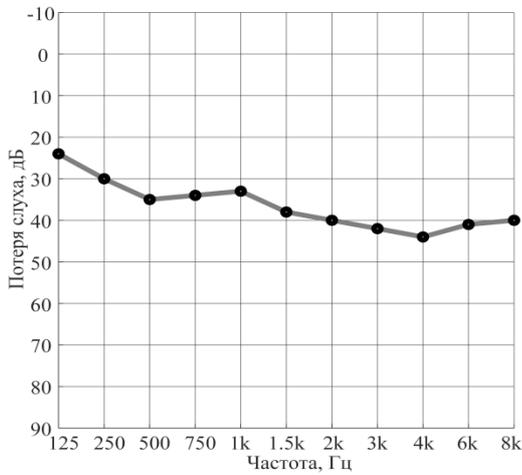


Рисунок 7 – Пример аудиограммы

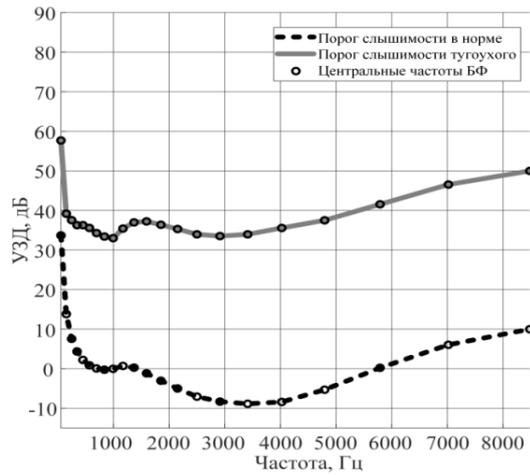


Рисунок 8 – Пороги слышимости

Поскольку дальнейшая обработка сигналов основывается на оценке уровня мощности в субполосе, то необходимо перейти от уровней звукового давления (УЗД) к «цифровым» дБ. Для этого требуется сместить порог слышимости в УЗД (рисунок 9) до уровня -90 дБ (рисунок 10).

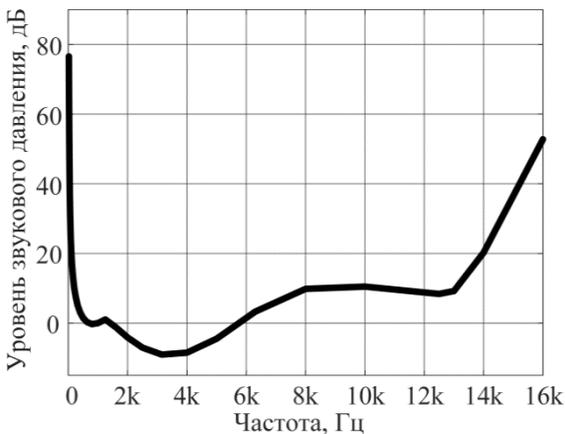


Рисунок 9 – Порог слышимости (УЗД дБ)

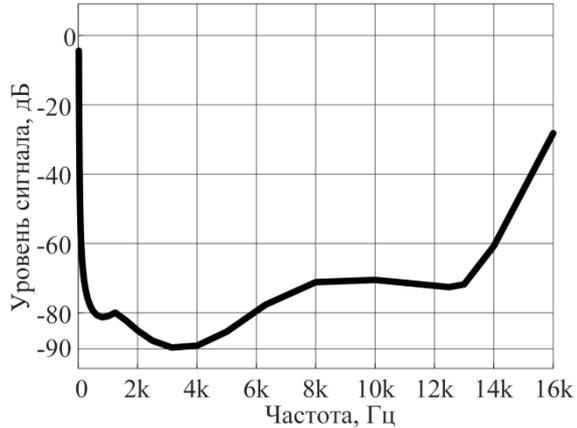


Рисунок 10 – Порог слышимости (цифр. дБ)

Чтобы настроить КДД на моделирование рекруитмента, необходимо знать его пороги (T_R) на центральных частотах банка фильтров. Для определения порогов рекруитмента используется надпороговая аудиометрия [6]. Однако далеко не всегда существует возможность получить эти данные. Решением проблемы может служить аппроксимация порогов рекруитмента по аудиограмме. На первом этапе вычисляется угол рекруитмента α (в градусах) [3]:

$$\alpha = 47 + 0.45HL(f_c) \quad (2)$$

Далее, зная HL и α , можно определить размер рекруитмента (R) на частоте f_c [3]:

$$R(f_c) = HL(f_c) / (\tan \alpha - 1) \quad (3)$$

Поскольку нас интересует именно порог рекруитмента (T_R) на частоте f_c , то по полученному размеру R , его можно вычислить следующим образом [3]:

$$T_R(f_c) = T_I(f_c) + R(f_c) \quad (4)$$

Стоит отметить, что при очень больших и очень малых уровнях потери слуха, величина R становится нереалистично маленькой. Чтобы избежать этой проблемы, предлагается установить минимально возможное значение параметра $R_{min} = 15$ дБ [3]. График зависимости размера рекруитмента от величины потери слуха

приведён на рисунке 11. Изменение угла рекруитмента в зависимости от уровня потери слуха отобразено на рисунке 12.

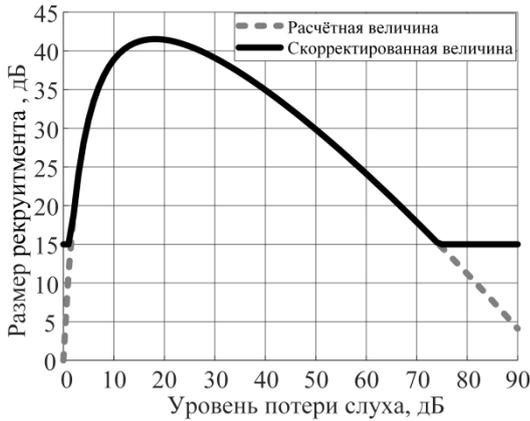


Рисунок 11 – Зависимость размера рекруитмента от уровня потери слуха

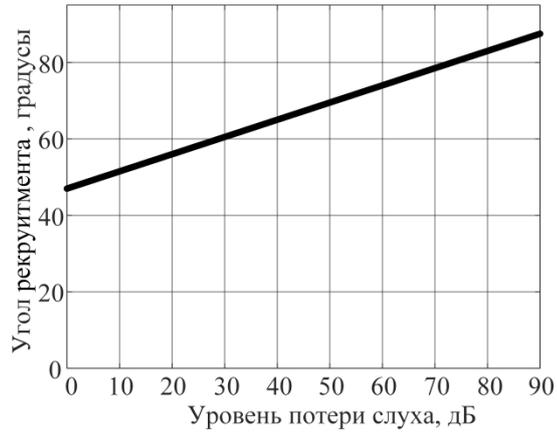


Рисунок 12 – Зависимость угла рекруитмента от уровня потери слуха

Получив из аудиограммы необходимые параметры слуха (T_I, T_N, T_R), можно настроить КДД.

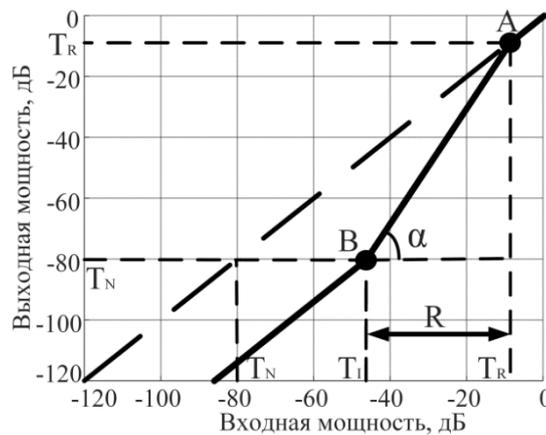


Рисунок 13 – Пример построения характеристики вход/выход КДД

На рисунке 13 приведён пример построения КДД для следующих значений: $T_N = -80.25$ дБ, $T_I = -46.6$ дБ, $T_{HL} = 33.65$ дБ, $\alpha = 62.14^\circ = 37.72$ дБ, $T_R = -8.87$ дБ. Штриховой линией под углом 45° изображена характеристика вход/выход КДД для нормально слышащего человека. На участке $T_R \dots 0$ дБ уровень входного сигнала повторяет уровень выходного, т.е. сигналы уровнем выше T_R дБ тугоухий человек воспринимает также как и нормально слышащий. На участке $T_I \dots T_R$ дБ громкость изменяется нелинейно (происходит моделирование рекруитмента и сужения слышимого динамического диапазона). На участке $-120 \dots T_I$ дБ сигнал ослабляется линейно, что позволяет не вносить нелинейных искажений в сигнал (в данном динамическом поддиапазоне).

Реализация КДД. В каждой обрабатываемой субполосе на выходе банка фильтров, на частотах, соответствующих центральным частотам фильтров, производится оценка мощности сигнала. С целью уменьшения вычислительной сложности расчёта мощности субполосного сигнала предлагается использовать экспоненциальное усреднение [8]:

$$P_k^{in}[n] = \begin{cases} (1 - AT)P_k[n - 1] + ATx_k^2[n], & \text{если } x_k^2[n] > P_k[n - 1] \\ (1 - RT)P_k[n - 1], & \text{если } x_k^2[n] \leq P_k[n - 1] \end{cases}, \quad k = 0 \dots K - 1, \quad (5)$$

где $P_k^{in}[n]$ – мощность сигнала на выходе k -го канала банка фильтров, $x_k[n]$ – сигнал на выходе k -го канала банка фильтров, K – число каналов банка фильтров, параметры AT и RT – коэффициенты атаки и восстановления [8]:

$$AT = 1 - \exp\left(\frac{-2.2T_S}{t_a/1000}\right), \quad RT = 1 - \exp\left(\frac{-2.2T_S}{t_r/1000}\right) \quad (6)$$

где T_S – период дискретизации, t_a и t_r – время атаки и восстановления, которые задаются в секундах.

Сама компрессия заключается в усилении/ослаблении субполосного сигнала [9]:

$$y_k[n] = x_k[n] \cdot g_k[n] \quad (7)$$

где $x_k[n]$ – сигнал на выходе k -го канала банка фильтров, $g_k[n]$ – коэффициент усиления/ослабления сигнала в k -й субполосе, $y_k[n]$ – выходной сигнал k -й субполосы.

Расчет субполосного коэффициента g_k ослабления выполняется в два этапа.

На первом этапе значения полученных оценок мощности в субполосах переводятся в дБ:

$$P_{k,dB}^{in} = 10 \log_{10} P_k^{in} \quad (8)$$

Затем выполняется расчёт уровней выходной мощности. Выходная мощность $P_k^{out}[n]$ рассчитывается исходя из характеристики вход/выход КДД.

На втором этапе вычисляются целевые коэффициенты ослабления в субполосах:

$$G_{k,dB} = P_{k,dB}^{out} - P_{k,dB}^{in} \quad (9)$$

Полученные коэффициенты переводятся из дБ в линейный масштаб:

$$g_k = 10^{G_{k,dB}/20} \quad (10)$$

Вывод. В работе изложены принципы построения КДД для моделирования нейросенсорной тугоухости по аудиограмме. Используя данный подход можно провести моделирование основных особенностей нейросенсорной тугоухости: сужения слышимого динамического диапазона и рекрутмента.

Литература

1. P. M. Zurek, J. G. Desloge Hearing loss and prosthesis simulation in audiology / The Hearing Journal, – 2007. – Vol. 60, Issue 7. – P. 32 – 33.
2. D. W. Kim et al. Simulation of hearing impairment with sensorineural hearing loss / Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Chicago, IL, USA, – 1997. – Vol. 5. – P. 1986 – 1989.
3. P. Duchnowski, P. M. Zurek Villchur revisited: another look at automatic gain control simulation of recruiting hearing loss / The Journal of the Acoustical Society of America, – 1995. – Vol. 98, Issue 6. – P. 3170 – 3181.
4. HeLPSV2 v2 User Guide / Sensimetrics Corporation, – Malden, Massachusetts, – 2014.
5. S. Banerjee The compression handbook: An overview of the characteristics and applications of compression amplification / Starkey Laboratories, – 2014. – 48 p.
6. А. Фонлантен, А. Хорст Слуховые аппараты / Феникс, Ростов н/Д. – 2009. – 304 с.
7. ISO 226:2003. Acoustics — Normal Equal-Loudness Level Contours / International Organization for Standardization, – Geneva, Switzerland, – 2003.
8. U. Zoelzer Digital audio signal processing / John Wiley & Sons Software, – 2008. – 334 p.
9. М. И. Порхун, М. И. Вашкевич Моделирование эффекта потери слуха / Цифровая обработка сигналов и ее применение: труды 20-й междунар. конф., Москва, 28-30 марта 2018 г., РНТОРЭС имени А.С. Попова, – Москва, 2018. – Т. 1. – С. 228 – 233.

SENSORINEURAL HEARING LOSS SIMULATION BASED ON THE SUBBAND DYNAMIC RANGE COMPRESSION

Porhun M.I., Vashkevich M.I.

Belarusian State University of Informatics and Radioelectronics
6, P. Brovki str., Computer Engineering Department, 220013, Minsk, Belarus,
e-mail: porhun@bsuir.by, vashkevich@bsuir.by

The basic principles of developing of dynamic range compressors for sensorineural hearing loss simulation by an audiogram are described. The model is based on the common using of subband decomposition of the audio signal and compression in each band. Loudness recruitment and narrowing of the audible dynamic range can be simulated using this approach. Dynamic range compressors are tuned based on parameters extracted from an audiogram of a person with sensorineural hearing loss.

