

## СПЕЦИФИКА НА ВЪЗПРИЕМАНЕТО НА ДИГИТАЛИЗИРАНА МУЗИКАЛНА ИНФОРМАЦИЯ

доц. д-р Алексей Стефанов  
гл. ас. д-р Румяна Жекова-Стойнова  
ВУ КТП и ЮЗУ „Неофит Рилски“, ПУ „П. Хилендарски“

## SPECIFICS OF PERCEPTION OF DIGITIZED MUSIC INFORMATION

Assoc. Prof. Aleksey Stefanov, PhD  
Chief. Ass. Romyana Zhekova-Stoynova, PhD

**Резюме:** Бурното развитие на цифровите технологии позволи в последните години рязко да се повиши равнището на цифровия звукозапис, на съхранението и възпроизвеждането на цифрови файлове, съдържащи различна по жанр и вид музикална информация. В доклада се представят различните видове цифрови звукови формати, степента им на достоверност и възможността за диференциране на отделните варианти в зависимост от специфичните им параметри.

**Ключови думи:** звук, праг на чуваемост, маскиране, mp3, загуба на качество, без загуба на качество.

**Abstract:** This paper will provide an overview of the fundamental principles of sound and the factors that determine our audio perception. The conversion of analog signals to digital form also known as „A/D conversion.“ has improved significantly the quality of audio recordings. One even says that high sampling rates and increased precision have erased any distinction between digital and analog. This paper is also an attempt to answer the question what is better - the lossless or lossy compression, the flac or mp3 format.

**Key words:** sound, hearing threshold, masking, mp3, flac, lossy, lossless.

### I. Възприемане на звука

Физиологичните особености на слуховия апарат на човека и психофизиологичните му характеристики имат определящо значение за рационалното проектиране и експлоатация на звукозаписваща и звуковъзпроизвеждаща апаратура.

Съгласуването на техническите средства за предаване на звука със субективните характеристики на възприятието способства да се достигне необходимата информационна достоверност на предаваните сигнали и да се осъществи в процеса на прослушване семантично и емоционално съответствие между първичния и възприемания звук образ.

Човешкото ухо е основен сетивен орган на аудиалната репрезентативна система. То представлява фин и деликатен, но същевременно много сложен и комплексен инструмент, който на практика извършва честотен и времеви анализ и реализира дискретно възприятие по честотен и динамичен диапазон.

Аналоговият сигнал, който постъпва на входа на слуховата система, се преобразува в последователност от електрически импулси, създавани от сложна

система, намираща се във вътрешното ухо. Впоследствие тази импулсна последователност се подава към съответните слухови анализатори в областта на големите подукълба. Като краен резултат се оформя звуков образ, който е адекватен на обективната звукова картина и съответства на нейната динамика във времето [1, 3].

Познаването в детайли на процеса на възникване на звуковите усещания позволява да се изградят психофизиологични модели, на базата на които се формират различни методи за обработка и кодиране на звуковата информация, намиращи изключително широко приложение в съвременните мултимедийни технологии.

Съществуват няколко основни момента [3]:

1. При човешката слухова система се разграничават три фази на предаване и трансформация на звуковата енергия:

- Разпространение в газова среда (външно ухо, слухов канал);
- Разпространение в твърда среда (средно ухо, костна система);
- Разпространение в течна среда (вътрешно ухо, ендолимфа).

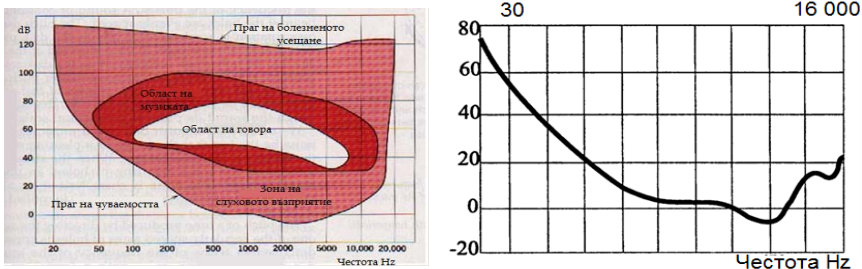
2. Човешкият слух честотно е ограничен в диапазона 16 – 20 000 Hz.

От гледна точка на възприемане на звукови нива също съществува долна и горна граница, които са показани на фиг. 1а [1]. Честотната зависимост на най-ниската възприемана гранична стойност на звуковото налягане, наречена „праг на чуваемост“, е представена на фиг. 1 б. Прагът на чуваемостта е различен за различните честоти. В областта на ниските и високите честоти звуковото налягане, което е необходимо, за да се появи усещане за звук, у човека е с по-високи стойности.

От диаграмите ясно се вижда наличието на честотна зависимост на предавателната характеристика на човешкият слухов орган, която е еволюционно изградена.

Формата на кривата на прага на чуваемостта се оформя до голяма степен от честотна филтрация, която се реализира във външното и средното ухо, като в този си вид определя най-добро възприемане на звукови колебания в диапазона, където е разположен честотният спектър на говора.

Стойностите на праговото възприемане се определят и от различните физиологични шумове в организма на човека, така че те да не са в състояние да възпрепятстват адекватното слухово възприятие.

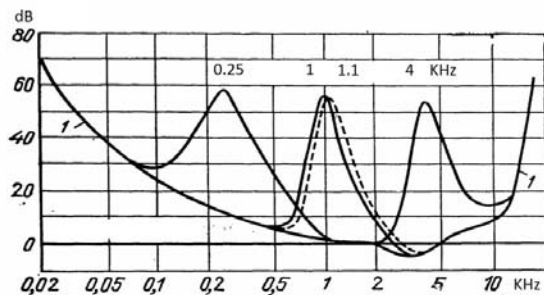


а)  
б)  
Фиг. 1. Нива на звуковото възприятие

3. Наличието на страничен източник на звук или шум изменя характера на зависимостта, илюстрирана на фиг. 1б [3]. Наблюдава се ефект на *маскиране* на полезния

сигнал, т.е. в областта, близо до честотата на смущаващия сигнал, съществено се повишава прагът на чуваемостта.

4. На фиг. 2 е представена честотната зависимост на прага на чуваемост при наличие на теснолентови смущаващи сигнали със средна честота 250, 1000, 1100 и 4000 Hz при ниво на смущаващия сигнал 60 dB.



Фиг. 2. Промяна на прага на чуваемостта при маскировка

При по-значителни нива на смущаващия сигнал полезният може въобще да не бъде чул. Явлението маскировка се проявява по различен начин за различни нива на смущаващия сигнал и при различни честоти. В посока към високите честоти спадът на кривата е по-полегат. Съществува маскиране по честота, по амплитуда и във времето [2, 4].

5. Друго много важно свойство на човешкия слух е неговата *нелинейност* [1]. При достатъчно високи нива на едночестотен тон човек започва да възприема и неговия втори, трети и т.н. хармоници. При прослушване на два тона се възприемат и сумарни, и разликови сигнали. Това свойство се използва при компресиране на звуковата картина, като се разчита на особеностите на човешкото възприятие за възстановяване на детайли от звуковата картина, които всъщност не биват предавани.

## II. Обработка, съхраняване и възпроизвеждане на звука

В природата звукът съществува под формата на механични колебания. За да бъде съхранен, пренесен и възпроизведен на друго място, различно от първоначалното, той обикновено се превръща в електрически сигнал. Още от времето на изобретяването на телефона е установено, че това е най-бързият и най-сигурният метод за обработка и пренос на звука.

Разбира се, съществуват и механични и оптични начини за запис и възпроизвеждане, но на практика те са свързани със сложно допълнително оборудване и не могат да осигурят многоцелевостта на електрическия сигнал, особено в дигиталния му вариант.

Масовото навлизане на компютърните технологии в ежедневието, развитието на цифровите комуникационни технологии решително определят водеща роля на цифровизиран звук пред аналоговия му вариант, въпреки все още нерешения категорично спор относно качеството, въздействието и близостта до първоизточника между двата варианта.

Аналоговият звуков сигнал може да бъде предаден на разстояние във вид на колебание на електрическото напрежение или на силата на тока [5]. В аудиоапаратура-та силата на тока или напрежението се променят със същата честота, с която се проме-

ния звуковият сигнал, а амплитудата на електрическия сигнал се изменя пропорционално на амплитудата на звуковата вълна.

Аналоговият сигнал с помощта на специални преобразувания може да бъде представен в цифров вид – последователност от някакви числа. Като следствие аналоговият звуков сигнал може да бъде въведен в компютър, обработен с цифрови методи и запазен на цифров носител във вид на набор от дискретни стойности.

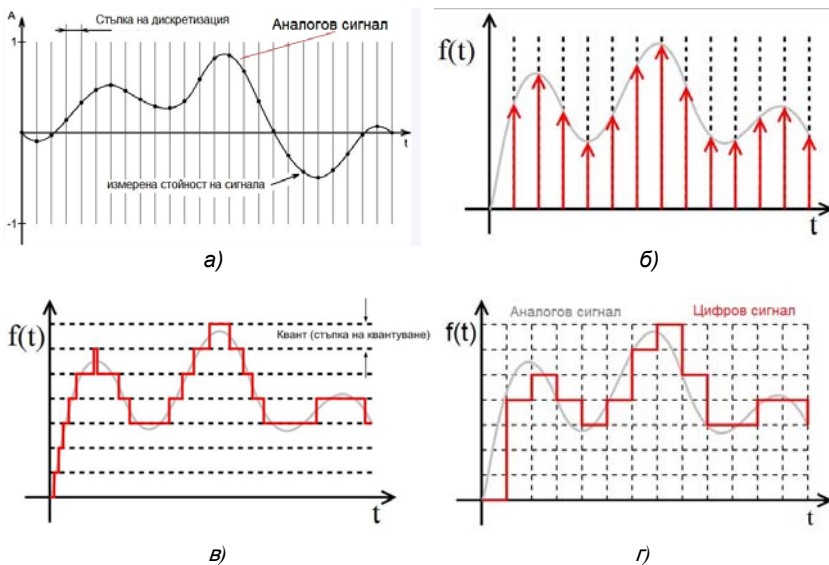
Важно е да се знае, че аналоговият или цифров звуков сигнал са само форма на представяне на звуковите колебания на материята, измислени от човека, за да може звукът да бъде анализиран и обработван. За да бъде чул цифровият сигнал, той трябва да се превърне отново в аналогов. А за да бъде чул аналоговият, е необходимо с негова помощ да бъде предизвикано трептене на някаква излъчваща повърхност, която и ще реализира вторично записаната звукова картина в околното пространство.

Графично изобразен, звуковият сигнал първоначално представлява постоянно променяща се, но непрекъсната линия (фиг. 3а) [5]. В този му вид обаче той не може да бъде съхранен на компютър.

Преобразуването на аналоговия сигнал в цифров вид се нарича аналого-цифрово преобразуване. Този процес преминава през две основни фази:

- Дискретизация – оценка на стойностите на амплитудата на аналоговия сигнал с определена стъпка във времето (фиг. 3б);
- Квантуване – запис на получената стойност на амплитудата в цифров вид (фиг. 3в).

Разликата между две съседни нива се нарича „стъпка на квантуване”, а разликата между дискретната стойност и присвоеното ниво – „грешка от квантуване” или още „квантов шум”. Стъпката на квантуване може да бъде както постоянна, така и променлива.



Фиг. 3. Преобразуване на аналоговия сигнал в цифров

Въпреки привидната простота и интуитивност на процеса на цифровизация на звука процедурата е свързана с редица проблеми.

Честотата на дискретизация (броят на измерванията в секунда) е определена теоретично като минимум два пъти по-голяма от най-високата честота, съдържаща се в обработвания звуков сигнал. На практика е необходимо тя да бъде избрана по-висока, за да може успешно да бъдат отстранени различни нежелани спектрални продукти. Колкото по-висока е тя, толкова по-малки са отрязъците от време, през които става отчитането на стойностите на амплитудата на сигнала, толкова по-близо е цифровизираният аудиосигнал до аналоговия първообраз.

Квантуването на стойностите на сигнала прибавя в спектъра му допълнителни смущения, наречени шум на квантуването. За да бъдат избегнати, се предприемат допълнителни мерки, които също усложняват процеса. И тук колкото повече са нивата на квантуване, толкова по-малка е грешката и по-голяма близостта до оригинала.

Въпреки съществуването на много на брой различни видове цифрови аудио файлове, като цяло всички съществуващи методи за кодиране на звукова информация могат условно да бъдат разделени на два типа [5, 6]:

1. Метод за компресиране на данни без загуби (англ. „lossless coding”) – това е метод за кодиране (уплътняване) на аудиоинформацията, позволяващ да се осигури стопроцентово възстановяване на изходните данни от компресирания поток. Под „изходни данни” тук се разбира изходният вид на цифровизираните аудиоданни. До този метод се прибегва тогава, когато е необходимо абсолютно, стопроцентово съхраняване на качеството на оригиналното звучене. Съществуващите съвременни алгоритми позволяват да се редуцира заеманият от данните обем от паметта с 20 – 50 % и при това да се гарантира безусловно адекватно възстановяване на изходните данни от компресирания материал. Кодирането без загуби, макар и идеално от гледна точка на запазването на аудиоматериалите, се оказва неспособно обаче да осигури високи нива на компресия и редукция на обема на цифровия поток.

2. Компресиране на данни със загуби (англ. „lossy coding”). Целта на такова кодиране е достигането на максимално висок коефициент на компресия на данните при съхраняване на качеството на звука на приемливи нива. В основата на този метод стоят две прости основополагащи съображения:

- Изходните цифрови аудиоданни притежават информационен излишък – те съдържат много ненужна за слуха информация, която може да бъде отстранена, като се увеличи по този начин коефициентът на компресия;
- Изискванията към качеството на звучене на аудиоматериала могат да бъдат различни и да зависят от конкретните цели и сфери на приложение.

Кодирането със загуби, както личи от името на метода, е свързано със загуба на някои части от аудиоинформацията. Такова кодиране като резултат довежда до това, че при възпроизвеждане звуковата картина много прилича на оригиналната, но на практика престава да бъде нейно оригинално отражение. В основата на множеството съществуващи начини за кодиране със загуби стои използването на психоакустичните особености на слуховата система на човека, а също така и различни хитрости, свързани с допълнителна техническа обработка на сигнала.

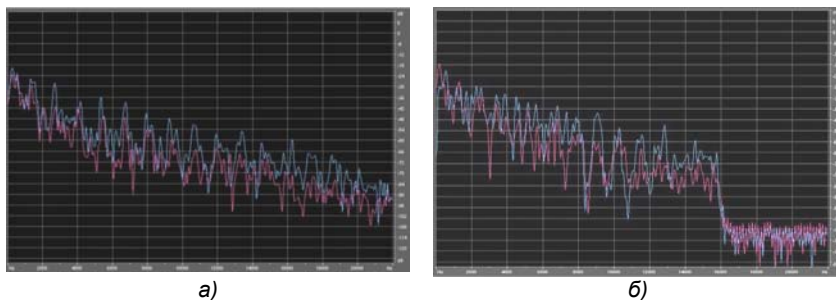
В процеса на кодиране данните се анализират за наличие на различни детайли и фрагменти, които могат да бъдат пренебрегнати. Маскирани честоти, нечуваеми и слабо чуваеми детайли – всичко това може да бъде пожертвано, за да се постигне по-висока стойност на коефициента на компресия. Аудиоинформацията в процеса на кодиране се подлага на значително опростяване там, където е важна преди всичко разбираемостта (в телефонията например честотите, по-високи от 4 КHz, не са необходими). Използването на подобен вид опростяване заедно с интелигентни квантуващи устройства

и удачно подбрани алгоритми за компресия позволяват да се постигне свиване на ползната информация спрямо първоначалната до 50 (петдесет) пъти.

Когато има по-високи изисквания към качеството на прослушване на аудиоматериалите (например портативни и битови аудио устройства), изходната звукова информация се подлага на по-щадящо кодиране. Трябва да се отбележи, че степента на агресивност на кодера по отношение на детайлите на звучене може да се регулира, като степента зависи от конкретната реализация. Съвременните кодери позволяват дори при сравнително висока степен на компресия (1:10) да се постигне отлично звучене, качеството на което средният слушател понякога определя като равно на това на изходните аудиоданни.

В последните години тече един непрекъснат спор, предимно в аудиофилските среди, относно качеството на записаната и възпроизведена след това музика във форматите със и без загуба.

На пръв поглед нещата не изглеждат толкова сложни. Ако сравним два файла с едно и също съдържание, но записани в различни формати (със загуби и без), е видно, че във файла със загуби са изрязани високите честоти над 16 KHz (фиг. 4а и б).



Фиг. 4. Честотна характеристика на музикален файл, компресиран без загуби и със загуби

В действителност, изхождайки от психофизиологичните особености на човешкия слух, във файла от фиг. 4б са изрязани и редица фрагменти от ниските и средните честоти (предимно в средните), които са маскирани или са с много ниско ниво, така че отиват под нивото на шумовете в тракта. В такъв случай не изглежда особено трудно разликата да бъде установена и слухово, а не само с апаратни средства. На практика това е възможно само при файлове с много голяма компресия и нисък битрейт до 250 kb/s (Побитовата скорост „*bit rate*” е скорост на обмен на информация. Използва се в контекста на степента на компресия на аудио- и видеофайловете и се означава с kb/s. Колкото по-висока е стойността, толкова по-добро е качеството.) При битрейт от 320 kb/s за файла със загуби да се установи разликата, е практически невъзможно.

Различни изследователи са правили много на брой тестове в различни условия и с различни групи от хора. Установено е, че когато тестът се извършва на сяло, като практика и провеждащият теста не е информиран за типа на прослушвания файл, между двата формата не са установени качествени различия.

Когато по някакъв начин обаче изпитваният има някаква предварителна информация, резултатите от теста лесно могат да бъдат манипулирани и да се получат предвидими резултати (ефект плацебо).

В крайна сметка аудиофилските среди продължават да твърдят, че разликата е ясно доловима за подготвения слух и не спират да инвестират във все по-скъпо оборудване, гонейки съвършения звук, докато любителите на музиката, които придават по-голяма тежест на самото съдържание, са напълно удовлетворени от качеството и на двата формата, без упорито да търсят разлика там, където субективно погледнато тя не съществува.

### Литература

1. Алдошина, И. и др. Электроакустика и звуковое вещание. Москва. Радио и связь, 2007.
2. Радзишевский, А. Основы аналогового и цифрового звука. Вильямс, Москва, 2006.
3. Blake, R., R. Sekuler. Perception. McGraw-Hill Higher Education, Singapore, 1990.
4. Chapman, N., J. Chapman. Digital Multimedia. John Wiley & Sons, Ltd. 2009.
5. Kirn, P. Real World Digital Audio. Peachpit Press, 2005.
6. Salomon, D., G. Motta. Handbook of Data Compression. Springer Verlag, London, 2010.