

Debreceni Egyetem
Informatika Kar

A pszichoakusztikai elfedés és szerepe a hangtömörítési eljárásokban

Témavezető:
Dr. Papp Zoltán
egyetemi adjunktus

Készítette:
Bertalan Éva
programtervező
matematikus

Debrecen
2007

Tartalomjegyzék

Bevezetés	3
1. A beszédfeldolgozás mennyiségei	5
1.1. Az inger mennyiségei	5
1.1.1. Hangnyomás szint (sound pressure level)	5
1.1.2. Hangintenzitás szint (sound intensity level)	6
1.1.3. Frekvencia (frequency)	6
1.1.4. Spektrum (spectrum)	6
1.1.5. Sűrűségi szint (density level)	7
1.2. Az érzet mennyiségei	7
1.2.1. Hangerősség (loudness level)	7
1.2.2. Hangosság (loudness)	8
1.2.3. Hangszínezet (tone)	8
2. Pszichoakusztikai alapfogalmak	10
2.1. Hallásküszöb	10
2.2. Kritikus sávszélesség	11
3. Elfedési minták	14
3.1. Tiszta hangok elfedése szélessávú zajjal	14
3.2. Tiszta hangok elfedése keskenysávú zajjal	15
3.3. Tiszta hangok elfedése aluláteresztő és feluláteresztő zajjal	17
3.4. Tiszta hangok elfedése tiszta hangokkal	17
3.5. Tiszta hangok elfedése összetett hangokkal	20
3.6. Pszichoakusztikai hangolási görbék	21
4. Időbeli hatások	25
4.1. Szimultán elfedés	26
4.2. Előelfedés	28
4.3. Utóelfedés	29
4.4. Időbeli elfedési minták	30
4.5. Elfedési periódus minták	37
4.6. Pulzációs küszöb	41
4.7. Kevert spektrális és temporális elfedés	45

TARTALOMJEGYZÉK	2
5. Az elfedés „kiegészítése”	46
5.1. A szimultán elfedés „kiegészítése”	46
5.2. Az utóelfedés „kiegészítése”	48
6. A szimultán elfedés pszichoakusztikai modellje	50
7. Hangtömörítés	52
7.1. Szokásos tömörítési eljárások	53
7.2. Az MPEG/audio tömörítés	55
7.3. Speex és Vorbis	60
7.3.1. Pszichoakusztikai alapötletek	60
7.3.2. A Vorbis modellje	61
Irodalomjegyzék	63

Bevezetés

Az akusztikai kommunikáció az emberi társadalom létezésének egyik alapfeltétele. Ezt a tényt figyelembe véve beláthatjuk, hogy az emberi hallórendszer fontos szerepet játszik. Hallórendszerünk azt a képességét, hogy az információkat befogadja, nemcsak a hang és a hatás közötti minőségi kapcsolat határozza meg, hanem az akusztikai inger és a hallási érzetek közötti mennyiségi kapcsolat is. Az új digitális audio technikák megjelenésével a hallórendszer akusztikai vevőként való értelmezésének tudománya, tehát a pszichoakusztika tudománya előtérbe került. Az emberi hallórendszer jellegzetességeit mindenképpen figyelembe kell venni akusztikai kommunikációs rendszerek tervezésénél és megvalósításánál.

Ez a dolgozat főként Eberhard Zwicker és Hugo Fastl *Psychoacoustics: facts and models* című könyvén alapszik. Az első fejezetben áttekintjük a későbbi tárgyalásokhoz szükséges mennyiségeket, ezekre szükség lesz ahhoz, hogy az elfedési jelenségekről beszéljünk. Az inger mennyiségeit már ismerhetjük a fizika világából, az érzet mennyiségei viszont más szemszögből jellemzik a hangokat. Az ezt követő fejezetben a hallásküszöbvel és a kritikus sávszélességgel foglalkozunk, amelyek szintén elengedhetetlenek az elfedés vizsgálatánál.

A harmadik fejezetben a zajokkal vagy más hangokkal elfedett tiszta hang jellemzőit írjuk le. Itt tárgyaljuk a pszichoakusztikai hangolási görbéket is, amelyek az elfedési jelenséget más nézőpontból írják le. Írunk az elfedés időbeli hatásairól is, vizsgáljuk a pulzációs küszöbvel összefüggő jelenségeket, majd végül az elfedés egy modelljét is ismertetjük. Az elfedés modelljét használják fel a veszteséges hangtömrítési eljárások.

Az elfedés fontos szerepet játszik a mindennapi életben. Egy csendes utcán való beszélgetés esetén például a beszélgetőknek alacsony hangerő is elég ahhoz, hogy megértsék egymást. Azonban ha elmegy mellettük egy hangos teherautó, akkor az a beszélgetést nagyon megzavarja. Ha a hangerőt nem változtatják, akkor nem fogják egymást hallani. Várhatnak, amíg a teherautó elmegy, hogy utána folytassák a beszélgetést, vagy beszélhetnek hangosabban. Ekkor ismét hallani fogják egymást. Hasonló jelenségekkel találkozhatunk a zenében is. Egy hangszer hangját elfedheti egy másik, ha az egyik hangos, míg a másik halk marad. Ha a hangos hangszer elhallgat, a halk hangszer hangja ismét hallhatóvá válik. Ezek tipikus példái a szimultán elfedésnek. Az elfedési jelenségeket rendszerint az elfedési küszöb meghatározásával mérjük. Az elfedési küszöb a teszthangnak (általában szinuszos teszt hang) az a hangnyomás szintje, amelyet egy elfedő hang jelenléte mellett éppen meghallunk. Az elfedési küszöböt - nagyon kevés speciális esettől eltekintve - az abszolút hallásküszöb felett találhatjuk; az elfedési küszöb azonos az abszolút hallásküszöbvel, ha az elfedő hang és a teszthang frekvenciája nagyon különbözik.

Ha az elfedő hang hangnyomása egyenletesen növekszik, az átmenet a hallható (fe-

detlen) teszthang és a teljesen elfedett teszthang között folyamatos. Ez azt jelenti, hogy a teljes elfedés mellett jelen van a részleges elfedés is. A részleges elfedés csökkenti a teszthang hangosságát, de nem fedi el teljesen. Ez a jelenség gyakran előfordul beszélgetésekben. A részleges elfedés jelenségével a dolgozat korlátozott terjedelme miatt nem térünk ki. Az elfedési jelenségeket nemcsak akkor mérhetjük, ha az elfedő hang és a teszthang egyszerre van jelen, hanem akkor is, ha egymás után szólalnak meg. Ez utóbbi esetben a teszthang egy rövid impulzus, amelyet az elfedő ingert bekapcsolása előtt hallunk. Ezt az elfedési jelenséget előelfedésnek nevezzük. Hatása nem túl erős, azonban ha a teszthang az elfedő hang kikapcsolása után jelentkezik, akkor már jelentős hatása van a hangra. Mivel ekkor a teszthang az elfedő hang befejezése után jelentkezik, ezt a jelenséget utóelfedésnek nevezzük.

Végül a hetedik fejezetben bemutatunk néhány tömörítési eljárást, majd megvizsgáljuk az MPEG és a Vorbis pszichoakusztikai modelljét. Ezek a tömörítési eljárások az elfedési jelenség sajátosságait is felhasználják.

1. fejezet

A beszédfeldolgozás mennyiségei

A hangot két szempontból tekinthetjük attól függően, hogy fizikailag vagy a pszichológiailag szeretnénk azt jellemezni. A fizikai mérés teljes mértékben objektív, a jellemzők jól meghatározhatóak. Ezt a hangot nevezzük ingernek, mivel ez váltja ki a fülben a hallást. A másik oldal teljesen szubjektív, azaz a mérés eredménye függ a vizsgált személytől. Ezeket az eredményeket általában átlagoljuk, és a hang pszichoakusztikai jellemzőinek, vagy érzetnek nevezzük. A két szempont szerinti mennyiségeket külön tárgyaljuk.

1.1. Az inger mennyiségei

1.1.1. Hangnyomás szint (sound pressure level)

A hang amplitúdója. A hallható hangot a közegben bekövetkező légnyomásváltozás hozza létre. A hangnyomás egysége pascal (Pa). A hang érzékelése során a hangnyomás időben változik. Ezt az időben változó mennyiséget jelöljük p -vel. A hangnyomást nem érzékeljük minden esetben. A legkisebb hallható érték az abszolút hallásküszöb, amit 10^{-15} Pa hangnyomáson észlelünk, a felső határ pedig a fájdalomküszöb 10^2 Pa hangnyomáson. Ezen két érték között érzékeljük a hangokat.

Beszédfeldolgozásban általában hangnyomás jelölésére a hangnyomás szintet (akusztikai decibel) használják. A szintet úgy vezették be, hogy az adott fizikai mennyiségnek kiválasztották egy alkalmas referencia értékét (a legkisebb még érzékelhető értéket), aztán vették a mennyiségek tízes alapú logaritmusát. A hangnyomás szint jele L , mértékegysége dB (decibel) és képlete:

$$L_p = 10 \lg \frac{p^2}{p_0^2} = 20 \lg \frac{p}{p_0} \text{ dB}, \quad (1.1)$$

ahol p_0 a referencia érték (az 1 kHz frekvenciájú, emberi fül által még éppen hallható hangnyomás), $p_0 = 20 \mu\text{Pa}$ a levegőben, $p_0 = 1 \mu\text{Pa}$ a vízben.

1.1.2. Hangintenzitás szint (sound intensity level)

Hangintenzitás alatt az egységnyi felületen időegység alatt átáramló hangenergiát értjük. Mértékegysége W/m^2 .

$$I = \frac{1}{T} \int_T^0 p(t) \cdot v(t) dt, \quad (1.2)$$

ahol p a hangnyomás, v a részecskék sebessége.

A hangintenzitás szint mértékegysége szintén dB, a hangintenzitásból a következőképpen származtatható:

$$L_i = 10 \lg \frac{I}{I_0} dB, \quad (1.3)$$

ahol $I_0 = 10^{-12} W/m^2$ szintén referenciaérték.

1.1.3. Frekvencia (frequency)

A fizikai mérések egyik legfontosabb mértékegysége a frekvencia. Ez a hanghullámokban egy másodperc alatti teljes változások számát fejezi ki, mértékegysége a Hertz (Hz). Tehát az 1 Hz azt jelenti, hogy egy esemény 1 másodperc alatt egyszer következik be.

Nagyon sok hangjelenség frekvenciafüggő. A hangokat nem a teljes frekvenciaskálán érzékeljük, hasonlóan a hangnyomáshoz, hanem egy jól meghatározható tartományban, ami körülbelül 20 Hz-től 20 kHz-ig terjed.

1.1.4. Spektrum (spectrum)

A hangokat jellemezhetjük időben frekvenciájuk függvényében. Azonban ez bonyolult és összetett adatokat eredményez, melyekkel nehéz dolgozni. Ugyanis a hangok a levegőben hullámok formájában haladnak, amely hullámok a szilárd akadályokon elnyelődhetnek, vagy visszaverődhetnek róluk. Amikor eljut a fülhöz, már egyenként időben eltérő hanghalmazt képez. Hogy ezt felbontassuk, szükség van egy olyan függvényre, amely invariáns az eltolásra, az összeadásra és a skalárral való szorzásra. Erre a problémára ad megoldást a Fourier-tétel, amely kimondja, hogy bármilyen $f(t)$ periódikus folyamat, amelynek periódusa T (azaz $T \neq 0$ az a legkisebb időköz, amelyre $f(t+T) = f(t)$), egyértelműen előállítható olyan, megfelelő amplitúdókkal és fázisállandókkal bíró harmonikus rezgések összegeként, amelyek körfrekvenciái az adott $f(t)$ rezgés körfrekvenciája ($\omega = 2\pi/T$) és ennek egész számú sokszorosai:

$$f(t) = c_0 + c_1 \sin(\omega t + \gamma_1) + c_2 \sin(2\omega t + \gamma_2) + \dots + c_n \sin(n\omega t + \gamma_n) + \dots, \quad (1.4)$$

ahol ω a körfrekvencia, c az amplitúdó, t az idő és γ a fázisállandó. Ez az úgynevezett Fourier-sorfejtés. A sor tagjait alkotó $c_n \sin(n\omega t + \gamma_n)$ harmonikus rezgéseket részrezgéseknek, ezek közül az ω körfrekvenciáját alaprezgésnek, a többi felharmonikusoknak nevezzük.

Ekkor tehát $X(f) = \frac{1}{\sqrt{2\pi}} \int x(t) e^{-i2\pi ft} dt$, ahol $x(t)$ a jel az idő függvényében. Ennek négyzete, $|X|^2$ az energiaspektrum, mely f -nek függvénye.

1.1.5. Sűrűségi szint (density level)

Ha valamilyen zajjal dolgozunk, akkor a hang intenzitása helyett érdekesebb a sűrűségét használni, azaz a hang 1 Hz sáv szélességén belüli intenzitását. A hang intenzitás és sűrűség logaritmikus összefüggését hang intenzitás sűrűségi szintnek, röviden sűrűségi szintnek nevezzük, jelölése l . Fehér zaj esetén, ahol a sűrűségi szint független a frekvenciától, l és L között a következő egyenlőség áll fenn:

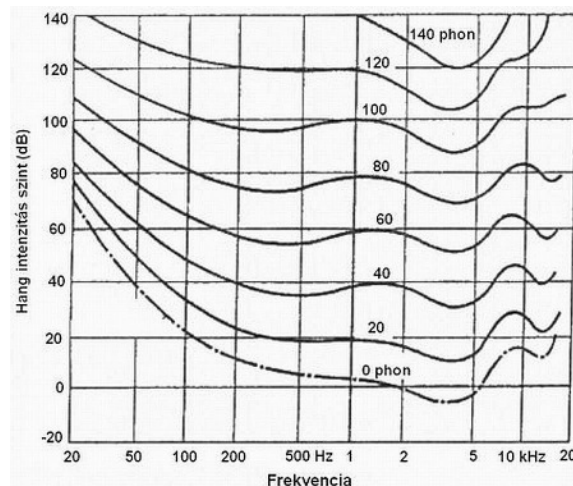
$$L = [l + 10 \lg(\Delta f / \text{Hz})] \text{ dB}, \quad (1.5)$$

ahol Δf a kérdéses hang sáv szélessége Hz-ben.

1.2. Az érzet mennyiségei

1.2.1. Hangerősség (loudness level)

A hangerősség jele L_n , mértékegysége phon. Egy hang hangerőssége annyi phon, ahány dB a vele azonos hangosságérzetet keltő 1 kHz-es szinuszos hang hangnyomásszintje. Ebből is látszik, hogy a pszichoakusztikai mértékegységek egyénekenként eltérnek, mivel mindenki máskor hall két hangot egyforma hangosnak. A 0 phon az a minimális érték, amit a fül érzékelni tud az adott frekvencián. Azonban ez a mértékegység nem felel meg a hangosság méréséhez, mert a phon értékek egyenletes változása nem vonja maga után a hangosságérzet egyenletes változását. Az érzet ugyanis nem logaritmikus, hanem hatványérzékelés. Ez a mértékegység viszont a dB miatt logaritmikus.



1.1. ábra. Fletcher - Munson görbék, vagy phon-görbék. A vízszintes tengely a frekvenciát mutatja Hz-ben, a függőleges a hangintenzitás szintet dB-ben. A görbék az azonos hangosságúnak érzékelt helyeket mutatják. A görbék felett a phon értékeket láthatjuk. A 0 phon-nal jelzett görbe jelöli az abszolút hallásküszöböt

1.2.2. Hangosság (loudness)

A hangok által kiváltott érzet összehasonlításán alapuló mérték, jele N , mértékegysége son. A hangosság egyenletes skálája a son-skála. Mivel ez már hatványon alapszik, megfelelő a hangosság érzet szerinti egyenletes méréséhez. Az értékek egyenletes változása a hangosságérzet egyenletes változását is jelenti. Megmutatja, hogy két phon-ban mért hangosságérzet aránya mekkora. Megállapították, hogy 10 phon hangosság szint növekedés a hangosságérzet kétszerezésének felel meg. (Kísérletileg a kétszerezést két azonos hangosságú hangforrás egyidejű bekapcsolásával valósíthatjuk meg, vagy az egy ill. kétfülű hallás közötti különbség felhasználásával.) Az átszámítási képlet a hangosság szint és a hangosság között:

$$\text{ha } L_n > 40 \text{ phon, akkor } N = 2^{\frac{L_n - 40}{10}} \text{ son.} \quad (1.6)$$

Ez azt jelenti, hogy például 40 phon = 1 son, 50 phon = 2 son, 60 phon = 4 son, ...

A son értékek összegezhethők (1 son + 1 son = 2 son), ha a hangok frekvenciáinak távolsága nagyobb, mint egy kritikus sáv. Ha a frekvenciák távolsága kisebb, mint egy kritikus sáv, akkor a fizikai intenzitások összegződnek.

Néhány hangforrás jellemzőit láthatjuk az 1.1. táblázatban.

Hangforrás	Hangnyomás	Hangnyomás szint	Hangosság
	pascal	dB	son
fájdalomküszöb	100	134	≈ 676
rövid ideig tartó hatás, ami halláskárosodást okoz	20	kb. 120	≈ 250
sugárhajtómű (100 m)	6 - 200	110 - 140	$\approx 125 - 1024$
légalapács (1 m), diszkó	2	kb. 100	≈ 60
hosszú ideig tartó hatás, ami halláskárosodást okoz	$6 \cdot 10^{-1}$	kb. 90	≈ 32
fűtő (10 m)	$2 \cdot 10^{-1} - 6 \cdot 10^{-1}$	80 - 90	$\approx 16 - 32$
vonat (10 m)	$2 \cdot 10^{-2} - 2 \cdot 10^{-1}$	60 - 80	$\approx 4 - 16$
tv készülék a lakásban (1 m)	$2 \cdot 10^{-2}$	kb. 60	≈ 4
normál beszélgetés (1 m)	$2 \cdot 10^{-3} - 2 \cdot 10^{-2}$	40 - 60	$\approx 1 - 4$
nagyon csendes szoba	$2 \cdot 10^{-4} - 6 \cdot 10^{-4}$	20 - 30	$\approx 0.15 - 0.4$
csendes légzés	$6 \cdot 10^{-5}$	10	≈ 0.02
hallásküszöb 2 kHz-en	$2 \cdot 10^{-5}$	0	0

1.1. táblázat. Hangforrások és jellemzőik

1.2.3. Hangszínezet (tone)

Különböző hangforrások egyenlő magasságú (alapfrekvenciájú) hangjai között általában „színezetbeli” különbséget fedezhetünk fel, így például a hegedű és a fuvola ugyanolyan magasságú hangjait meg tudjuk különböztetni egymástól. Helmholtz megállapítása szerint a hangszínezetet az alaphanghoz csatlakozó felhangok frekvenciája és viszonylagos erőssége,

azaz a hang spektruma szabja meg. Ha tehát a vizsgált zenei hangot a spektrumnál leírt módon Fourier-sorral állítjuk elő, a hangszínezet a rezgések c_1 , c_2 , ... amplitúdóinak arányától függ.

A hangszínezetet meghatározó hangspektrum a hangforrás természetén kívül a hangkeltés módjától is függ. Aszerint, hogy a felhangok közül a mélyebbek vagy a magasabbak uralkodnak, a hangot tompának vagy élesnek érzékeljük.

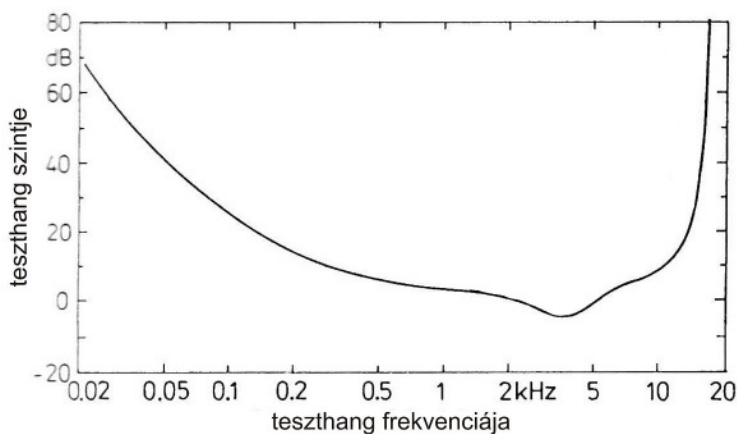
2. fejezet

Pszichoakusztikai alapfogalmak

2.1. Hallásküszöb

Környezetünkben a teljes csend állapota sosem tapasztalható meg. Amit mi teljes csendként érzékelünk, az igazából többféle hangból tevődik össze: a levegőmolekulák mozgása, az izmaink mozgása, és sok egyéb hasonló zaj hallható ilyenkor. Ezek a zajok túlságosan nagy terhelést jelentenének az agy számára, nem tudja őket folyamatosan feldolgozni. Ezért alakult ki a hallásküszöb, ami megvédi az agyat a felesleges információ-tömegtől.

Figyeljük meg az egyes frekvenciákon, hogy hány dB hangnyomással rendelkezik az a tiszta hang, amit még éppen hallunk. A teljes frekvenciatartományban ez fogja szolgáltatni az abszolút hallásküszöböt, ami tulajdonképpen a természetes zajok speciális elfedési görbéje. A hallásküszöb alakja a következő:

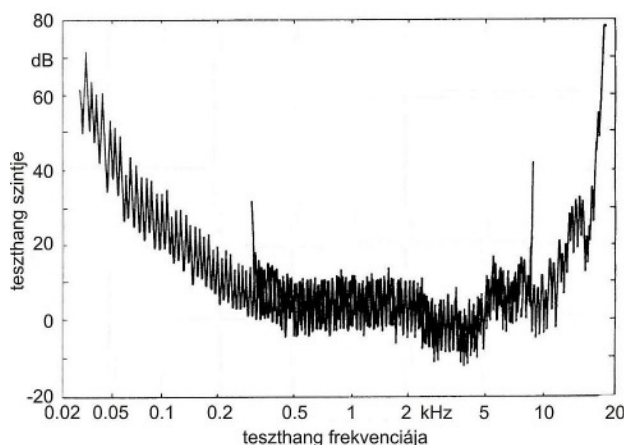


2.1. ábra. Az abszolút hallásküszöb, vagyis a természetes alapzörejek elfedési görbéje

Ezt a görbét könnyen meghatározhatjuk például a Békésy - féle követési módszerrel. A módszer a következőképpen működik: a vizsgált személy egy visszhangmentes és zajmentes szobában szemből érkező tiszta szinuszos hangot hall. Van egy kapcsolója, amivel a hangnyomást tudja csökkenteni vagy

növelni. A tiszta hang frekvenciája a vizsgálat során viszonylag lassan vándorol az alacsonytól a magas frekvencia felé. Az alany a kapcsolót úgy használhatja, hogy amikor a hangot határozottan hallja, akkor a kapcsoló átállításával a hangnyomást csökkenő állapotba viszi át, amikor pedig a hang már határozottan nem hallható, akkor a kapcsolót visszaállítja, így a hangnyomás ismét növekedni kezd.

Ekkor a következő függvényt kapjuk:



2.2. ábra. Az abszolút hallásküszöb meghatározása a Békésy-féle követési módszerrel. Az ábra a teszthang éppen érzékelhető szintjét mutatja frekvenciájának függvényében. A hallásküszöböt 0.8 kHz és 8 kHz között kétszer mértük

A középvonal fogja adni a hallásküszöböt.

A hallásküszöb minden ember esetében máshogy helyezkedhet el. A normál görbét, ami a 2.1. ábrán is látható, az olyan 20-25 éves embereknél mérték, akiknek nincs halláskárosodásuk. A görbe minimuma 1 kHz és 5 kHz között található, hallásunk itt a legérzékenyebb, ami egyúttal azt is jelenti, hogy itt sérülhet meg a legkönnyebben. A halláskárosodás származhat hangos zene hallgatásától fülhallgatón keresztül, de akár a zajos gépek közelében való munkavégzéstől is. Ekkor a küszöb eltolódik a magasabb hangnyomás-értékek felé.

A hangokat frekvenciatartományuk alapján három részre szoktuk osztani:

infrahang	20 Hz alatt
hallható hang	20 Hz - 20 kHz
ultrahang	20 kHz felett

2.2. Kritikus sáv szélesség

A kritikus sávok ötlete a fül azon tulajdonságából adódott, hogy a hallható frekvenciatartományt alsávokra osztva analizálja. Több egyidejű komponens összegződése eltérő módon történik, ha azok egy bizonyos távolságon belül vagy kívül vannak. Ezeket a távolságokat nevezzük kritikus sáv szélességeknek (critical bandwidth). A fogalmat Fletcher vezette be 1940-ben.

Egy keskenysávú zaj sávszélességét egy ideig növelhetjük anélkül, hogy a hallási szabály megváltozna. Egy bizonyos értéken felül viszont teljesen más szabályok lépnek érvénybe. Hangosságérzetünk hasonlóképpen viselkedik. Akár sok szinuszos hangról, akár zajról van szó, a sáv szélesedésével csak a hangintenzitás változása szerint halljuk hangosabbnak a jelet. Ha viszont a kritikus sávszélességet átlépjük, akkor a hangosság érzete erőteljesebben növekszik, mint a hang intenzitása, ekkor már a son értékek fognak összeadódni. A hallástartomány 24 összeérő kritikus sávval fedhető le, ezek frekvenciahatárait és középfrekvenciáit táblázatba foglalták. Két hang egy sávba tartozását mindig úgy kell megállapítani, hogy az adott két hang által meghatározott frekvenciaköz kisebb-e vagy nagyobb, mint amekkora a táblázati szomszédos kritikus sávok értéke. A 2.1. táblázat értékeit különböző kísérletek eredményei adták.

Sávszám (Bark)	Középfrekvencia (Hz)	Alsó-felső határ (Hz)	Sávszélesség (Hz)
1	50	0 - 100	100
2	150	100 - 200	100
3	250	200 - 300	100
4	350	300 - 400	100
5	450	400 - 510	110
6	570	510 - 630	120
7	700	630 - 770	140
8	840	770 - 920	150
9	1000	920 - 1080	160
10	1170	1080 - 1270	190
11	1370	1270 - 1480	210
12	1600	1480 - 1720	240
13	1850	1720 - 2000	280
14	2150	2000 - 2320	320
15	2500	2320 - 2700	380
16	2900	2700 - 3150	450
17	3400	3150 - 3700	550
18	4000	3700 - 4400	700
19	4800	4400 - 5300	900
20	5800	5300 - 6400	1100
21	7000	6400 - 7700	1300
22	8500	7700 - 9500	1800
23	10500	9500 - 12000	2500
24	13500	12000 - 15500	3500

2.1. táblázat. *Kritikus sávok jellemzői*

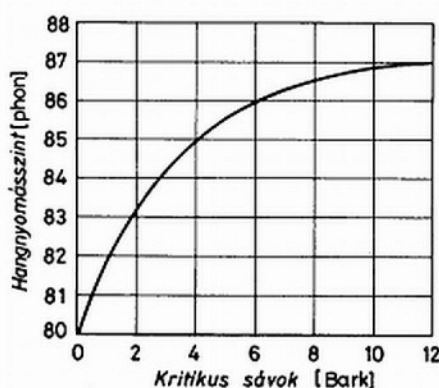
A kritikus sávok tehát 500 Hz alatt 100 Hz szélesek, 500 Hz felett pedig a sávszélesség a kritikus sáv középfrekvenciájától függ (500 Hz és 3500 Hz között körülbelül kisterc, 3500 Hz felett körülbelül nagyterc). Külön érdekesség az, hogy míg a hallásküszöb mindenkinél

egyedi, a kritikus sávok majdnem teljesen függetlenek attól, hogy ki a vizsgált személy.

Ott célszerű a kritikus sávoknál keskenyebb sávokkal dolgozni, ahol elfedési jelenség van, így az elfedési görbék teljesítményben összegezhetők.

A Bark-skála két frekvencia különbségét pszichoakusztikai szempontból megmutató skála. A különbség jellemzője, hogy hányszor lehet felmérni a tartományban a kritikus sáv szélességet. Például a 150 - 1000 Hz-es frekvenciatartományban 8 kritikus sáv van, ez tehát 8 Bark távolság.

Példaképpen nézzük meg, hogy két 77 phon-os hangforrás eredő hangosság szintje hogyan változik, ha növeljük a közöttük lévő frekvenciakülönbséget (2.3. ábra). Kritikus sávon belül (0 Bark) az eredő 3 phon-nal növekszik, így lesz 80 phon. Az emelkedés eleinte gyors, majd lelassul, s végül eléri a 87 phon eredőt, azaz a 10 phon növekedést. Ez utóbbi esetben a son-szabályt alkalmazzuk, a számítás tehát a következő lesz: 77 phon = 13 son, mivel kritikus sávon kívül a son értékek összegződnek, ezért a két hang együttesen $2 \cdot 13$ son = 26 son, ami átszámítva 87 phon.



2.3. ábra. Két 77 phon-os hangforrás eredő hangosság szintje a frekvenciák távolságának függvényében

A kritikus sávoknak fontos szerepük van még a kéthang (dichotikus) hallás vizsgálatában, ahol két hangot akkor hallunk különbözőnek, ha azok távolsága nagyobb, mint az adott frekvenciához tartozó kritikus sáv szélesség. Másik alkalmazási területe két hangköz távolságainak vizsgálata. Két hangközt akkor érzünk azonos távolságúnak, ha a hangközöket alkotó két hang távolsága egy kritikus sáv. Ez a jelenség független a kritikus sáv középfrekvenciájától. Például egy 3500 Hz feletti nagytercet és egy 100 Hz környékén elhelyezkedő oktáv hangközt azonos távolságúnak érzünk.

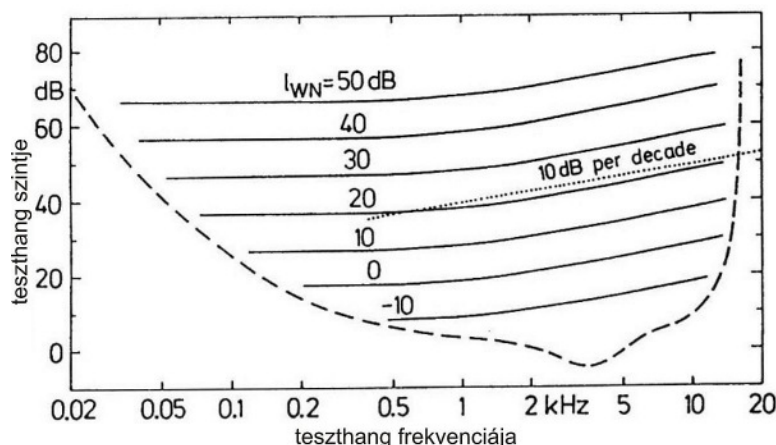
3. fejezet

Elfedési minták

3.1. Tiszta hangok elfedése szélessávú zajjal

A legkönnyebben definiálható szélessávú zaj a fehér zaj (white noise), mivel spektrális sűrűsége független a frekvenciától. A szélessáv terjedelme a vizsgálatokban definíció szerint 20 Hz-től 20 kHz-ig tart, ebben a tartományban érzékeljük ugyanis a hangokat.

A 3.1. ábra mutatja azokat a küszöböket, melyeket a teszhang frekvenciájának függvényében ábrázolunk a fehér zaj különböző sűrűségi szintjei mellett. Az ábrán látható még a hallásküszöb is szaggatott vonallal jelölve.



3.1. ábra. Adott szintű (l_{WN}) fehér zajok által elfedett teszhang szintje a teszhang frekvenciájának függvényében. A szaggatott vonal az abszolút hallásküszöböt jelöli

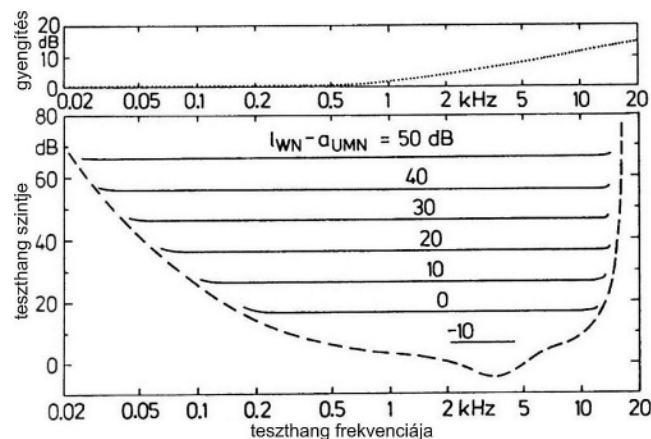
Az elfedés alkalmazása esetén kapott küszöbérték alacsony frekvenciákon 17 dB-el a fehér zaj adott sűrűségi szintje fölött található. Ez azt jelenti, hogy a zaj és a tiszta hang együtt 17 dB-el magasabb szinten hallható meg, mint a zaj a teszhang nélkül. Érdekes jelenség az, hogy még a negatív sűrűségi szintek is produkálnak elfedést.

A görbét 10 dB-el fentebb tolva megkapjuk a 10 dB-el magasabb szintű zaj elfedési görbéjét, ami a szélessávú zajjal történő elfedés lineáris tulajdonságát mutatja.

Nagyon alacsony és nagyon magas frekvenciákon az elfedés értéke megegyezik a hallásküszöb értékével. Meglepő módon a jelentős frekvenciabeli eltérések a hallásküszöb

majdnem teljesen eltűnnek, ha szélessávú zajjal fedjük el a hangot.

Az ábrából látható, hogy bár a fehér zaj spektrális sűrűsége teljesen frekvenciafüggetlen, az elfedési görbék, amelyeket a tiszta hang és a fehér zaj együttes megszólaltatásával kapunk, már függnek bizonyos mértékben a frekvenciától. Ezek az értékek folytonos vonallal vannak jelölve. Frekvenciafüggetlenséget egyedül az alacsony frekvenciákon mutatnak, körülbelül 500 Hz-ig. 500 Hz felett azonban már fokozatosan emelkednek az elfedési görbék. Ez az emelkedés a frekvenciával arányos, mértéke tízszeres frekvencia esetén körülbelül 10 dB.



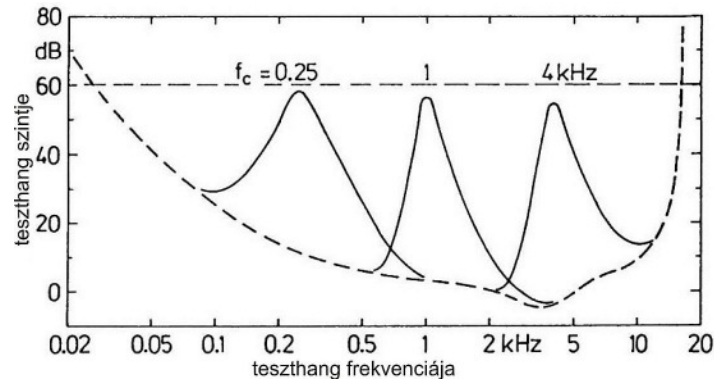
3.2. ábra. Adott szintű egyenletes elfedési zajok (UMN) által elfedett teszthang szintje a teszthang frekvenciájának függvényében. A felső rajz (pontozott vonal) azt a gyengítést mutatja, amely ahhoz szükséges, hogy fehér zajból egyenletes elfedési zajt állítsunk elő

Ezek az elfedések függetlenné tehetők a frekvenciától a teljes frekvenciaskálán egy speciális zaj segítségével, amely megfelelő módon függ a frekvenciától. Ezt a zajt egyenletes elfedési zajnak (uniform masking noise) nevezzük, mivel egy olyan elfedést állít elő, ami független a frekvenciától. A 3.2. ábrán látható, milyen gyengítést kell alkalmazni a fehér zajra, hogy megkapjuk ezt a speciális zajt, az ábra alsó részén pedig látható az egyenletes elfedési zaj alkalmazásának eredménye. Jól megfigyelhető, hogy az így kapott elfedési görbék már egyenletesek minden frekvencián, azaz függetlenek attól.

3.2. Tiszta hangok elfedése keskenysávú zajjal

A keskenysávú zaj jelen esetben olyan zajt jelent, aminek sáv szélessége egyenlő vagy kisebb, mint a kritikus sáv szélesség. Mint ahogy a 2.2. szakaszban tárgyaltuk, ez azért kényelmes, mert az energiákat ekkor összegezhethetjük, azonban ha a sáv szélessége meghaladná a kritikus sávét, akkor speciális szabályokat kellene alkalmazni. A zaj sűrűségi szintje helyett most használjuk a teljes szintjét, így az adatok több dolgot árulnak el. Az 1.5. egyenlőséget felhasználva a sűrűségi szintet könnyű áttanszformálni, ha a sáv szélesség adott.

A 3.3. ábrán három elfedési görbe látható. A keskenysávú zajok szélessége a kritikus sáv szélesség, középfrekvenciájuk 0.25, 1 és 4 kHz, szintjük egységesen 60 dB. A sáv szé-

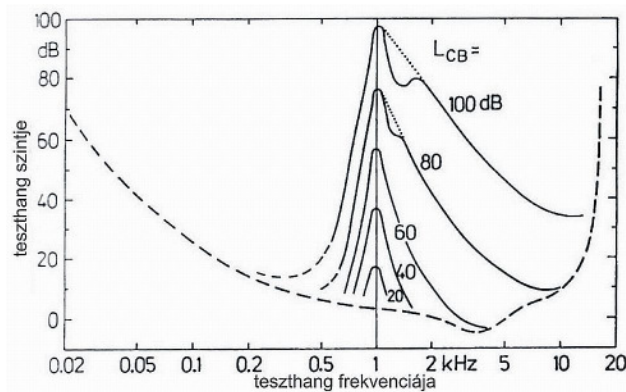


3.3. ábra. 60 dB szintű kritikus sáv szélességű zajok által elfedett teszthang szintje. A zaj középfrekvenciái 0.25, 1 és 4 kHz. A szaggatott vonal ismét az abszolút hallásküszöböt jelöli

lességet könnyen meghatározhatjuk a 2.1. táblázatból. A 0.25 kHz-hez így 100, 1 kHz-hez 160, 4 kHz-hez 700 Hz sáv szélesség tartozik.

A görbék elemzése rögtön feltűnik, hogy azok eléggé meredek, esetenként több mint 200 dB/oktáv a meredekségük. Az is feltűnő, hogy az 1 kHz és 4 kHz középfrekvenciájú elfedés eredménye nagyon hasonló, azonban a 250 Hz középfrekvencia esetében a görbe szélesebb, meredeksége kisebb.

Szintén eltér a görbék maximumának távolsága a zajok szintjétől, ami 60 dB minden esetben, és amit az ábrán a vízszintes szaggatott vonal jelöl. A távolság 250 Hz középfrekvencia esetén 2 dB, 1 kHz középfrekvencia esetén 3 dB és 4 kHz középfrekvencia esetén 5 dB. A görbék meredeksége a középfrekvenciáktól alacsonyabb frekvenciákon nagyobb, és valamivel kisebb a magasabb frekvenciák mellett.



3.4. ábra. 1 kHz-es középfrekvenciájú és különböző szintű kritikus sáv szélességű zajok által elfedett teszthang szintje a teszthang frekvenciájának függvényében

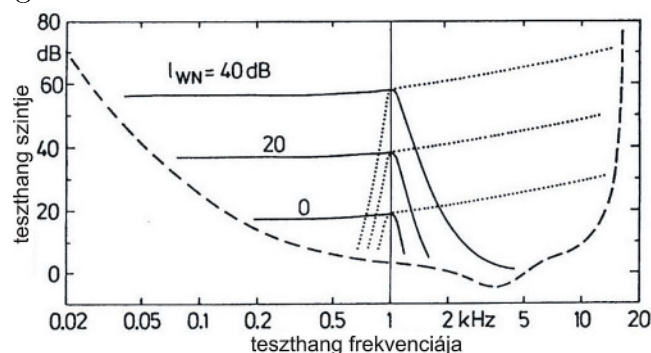
Ez tehát a görbék eltérése, ha azonos szinten, de más középfrekvenciákon vizsgáljuk őket. Most nézzük meg az 1 kHz-es középfrekvenciájú keskenysávú zajt, hogyan viselkedik, ha a szintjét változtatjuk. Ezeket az eredményeket a 3.4. ábra mutatja. Minden elfedési görbe nagyon meredek emelkedést mutat a maximuma előtt, ezen a szakaszon ráadásul párhuzamosak is. A görbék maximuma minden esetben a zaj szintjétől 3 dB-el alacso-

nyabban található, tehát ez a tulajdonság független a zaj szintjétől.

A görbe a maximuma után gyorsan csökkenni kezd. Alacsony és közepes zajszinteken ez a csökkenés folytonos és meredek, azonban magasabb szintek felé haladva a meredekség egyre kisebb. Ezért megállapíthatjuk, hogy az elfedési görbék frekvenciafüggősége függ a szinttől, vagyis nemlineáris a közöttük lévő kapcsolat. A magasabb zajszinteken ráadásul még egy másik jelenséggel is számolni kell, ezt szemlélteti az ábra 80 és 100 dB-es zajszinteken. A lebegések ezeken a szinteken mind az elfedő, mind a teszthang megfigyelését lehetetlenné teszik, helyettük egy másfajta zajt hallunk, amelyet a teszthang és a keskenysávú zaj közötti kölcsönhatás hoz létre. A teszthang szintjének csökkenésével ez a jelenség eltűnik, újra hallhatjuk az elfedett teszthangot. A lebegési jelenségnél a teszthang csak akkor válik hallhatóvá, ha a szintjét a pontozott vonallal jelölt értékre emeljük.

3.3. Tiszta hangok elfedése aluláteresztő és felüláteresztő zajjal

Ebben az esetben a használt zaj fehér zaj, amelyet egy meredek aluláteresztő szűrővel, valamint egy nagyon meredek felüláteresztő szűrővel korlátozunk. A vágási frekvencia az első esetben 1.1 kHz, a másodikban 0.9 kHz. Az elfedési görbék a 3.5. ábrán láthatóak, a paraméter a sűrűségi szint.



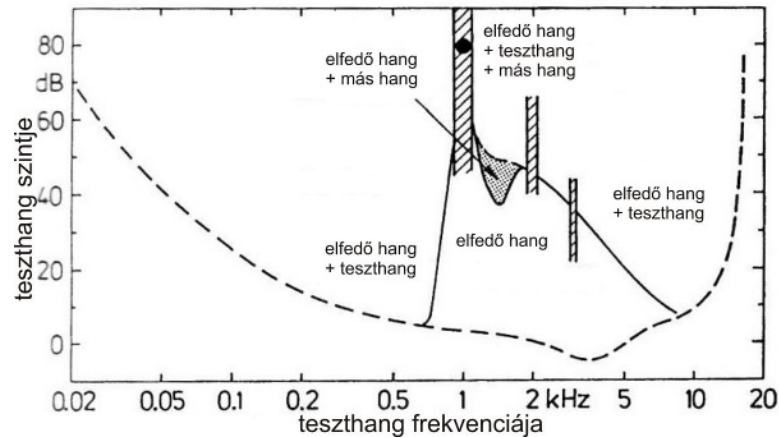
3.5. ábra. Különböző szintű aluláteresztő (folytonos görbék) és felüláteresztő (pontosított görbék) zajok által elfedett teszthang szintje a teszthang frekvenciájának függvényeként. A felüláteresztő zaj vágási frekvenciája 0.9 kHz, az aluláteresztő zaj vágási frekvenciája pedig 1.1 kHz

A vágási frekvencián a görbék nem a zaj gyengítésének meredekségével csökkennek, hanem a keskenysávú zajnál bemutatott módon. Az aluláteresztő zaj vágási frekvenciája alatt az elfedés megegyezik a fehér zaj esetén látottakkal. Ugyanez igaz a felüláteresztő zaj esetén a vágási frekvencia felett. Látható a fehér zajjal való elfedésre jellemző emelkedés is, amelynek mértéke tízszeres frekvencia esetén körülbelül 10 dB.

3.4. Tiszta hangok elfedése tiszta hangokkal

A következő elfedés fajta sok nehézséget okoz, főleg közepes és magas zajszinteken. A 3.6. ábra egy tiszta hang elfedését mutatja. A tiszta hangot frekvenciájának függvényeként

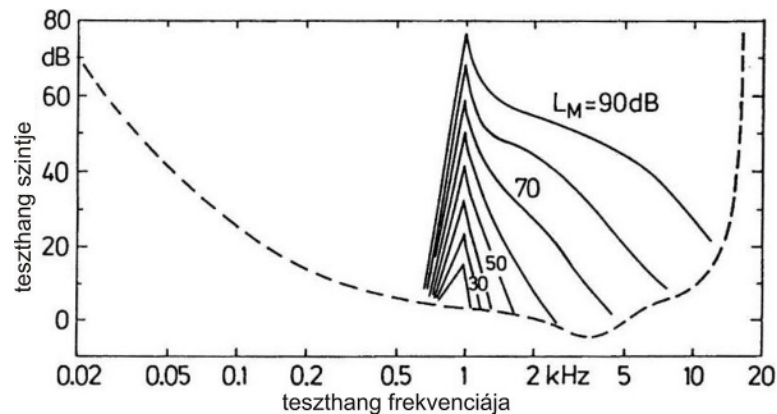
ábrázolja, az elfedő hang frekvenciája 1 kHz, szintje 80 dB.



3.6. ábra. 1 kHz frekvenciájú 80 dB szintű elfedő hang által elfedett teszthang szintje a teszthang frekvenciájának függvényében. Az egyes területeket különböző érzetek jellemzik. A keresztben vonalkázott területekre például jellemzőek a lebegések

Az első jelentős különbség az eddig tárgyaltakhoz képest az, hogy megjelennek a lebegések, ha a teszthang frekvenciája az 1 kHz-es elfedő hang közelében van. Az egész frekvenciatartományt megfigyelve még két ilyen esetet találunk, az egyiket 2 kHz, a másikat 3 kHz környékén.

A másik problémás terület a teszthang 1.4 kHz frekvenciájú környéke. A teszthang viszonylag alacsony, 40 dB-es szintje mellett megjelenik egy újabb hang. Ebben a tartományban gyakran nem hallani a teszthangot, helyette ezt az újabb hangot lehet felismerni 600 Hz közelében. Küszöbe nem a teszthang küszöbe lesz, a teszthang csak a körülbelül 50 dB feletti szinteken ismerhető fel.



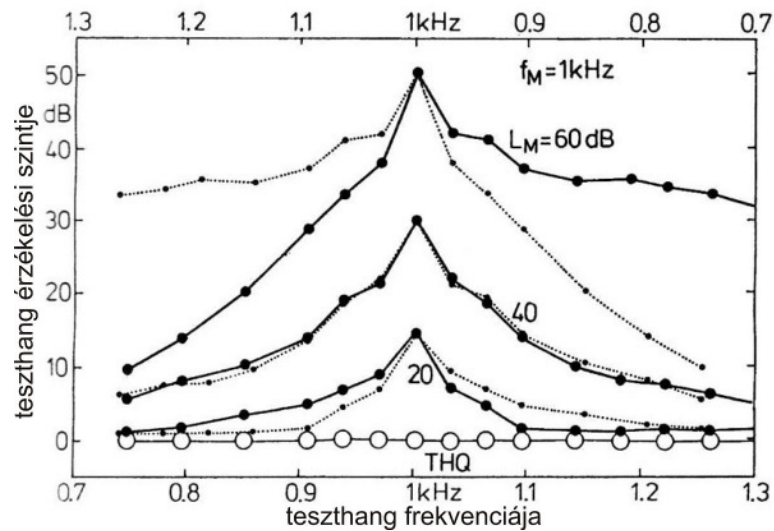
3.7. ábra. 1 kHz-es, különböző szintű hangok által elfedett teszthang szintje a teszthang frekvenciájának függvényében. A görbék alakját 1 kHz környékén csak becsülni tudjuk

Az ábrán az elfedés különböző területei vannak ábrázolva. A teszthang hallásküszöbe alatt csak az elfedő hang hallható. Körülbelül 700 Hz alatt, ha a teszthang hangnyomás szintjét a hallásküszöb fölé emeljük, akkor egy olyan területet láthatunk, ahol a teszthang

és az elfedő hang együtt hallható. Körülbelül 700 Hz és 9 kHz között egy olyan régió található, ahol csak az elfedő hang hallható annak ellenére, hogy a teszthang küszöbe az elfedő hang nélkül sokkal alacsonyabb. Azok a területek, ahol a lebegések hallhatóak, vonalkázással vannak jelölve. Pontozással jelöltük azt a helyet, ahol csak az elfedő hang és az új hang hallható (a teszthang nem). A teszthang elfedési küszöbe fölött, 1 kHz és 2 kHz között ugyancsak ez az új hang hallható az elfedő hang és a teszthang mellett.

Az új hang megfelelő eszközökkel méréselkelhető, ezért viszonylag jól lehet mérni az elfedett teszthangot a zavaró tényező ellenére is. A lebegések területét azonban nem lehet kikerülni. Ekkor viszont rendelkezünk egy mérhető ponttal, amely ott található, ahol a teszthang és az elfedő hang frekvenciája azonos. Ezt mutatja az ábrán a pont. A 3.7. ábra ennek a módszernek az eredményét ábrázolja. Látható, hogy a meredekség egyre kisebb az alacsonyabb frekvenciák felé az elfedő hang szintjének csökkenésével. Az 1 kHz-nél nagyobb értékeken azonban a szint növekedésével az elfedési görbék ellaposodnak.

Az elfedési görbe alacsony és magas frekvencián látható meredekségeinek különbsége alacsony hangnyomás szinteken meglepő tulajdonsággal rendelkezik. Itt ugyanis az elfedési görbe laposabb alacsonyabb frekvenciákon, mint magasabb frekvenciákon. Magasabb szinteken azonban ez fordítva teljesül. A görbe itt a magasabb frekvenciákon lesz laposabb. A magasabb szintek mellett tapasztalható jelenség már ismerős a keskenysávú zajokkal történő elfedésből. A két jelenség határa 40 dB környékén található, itt a görbe majdhogynem teljesen szimmetrikus.



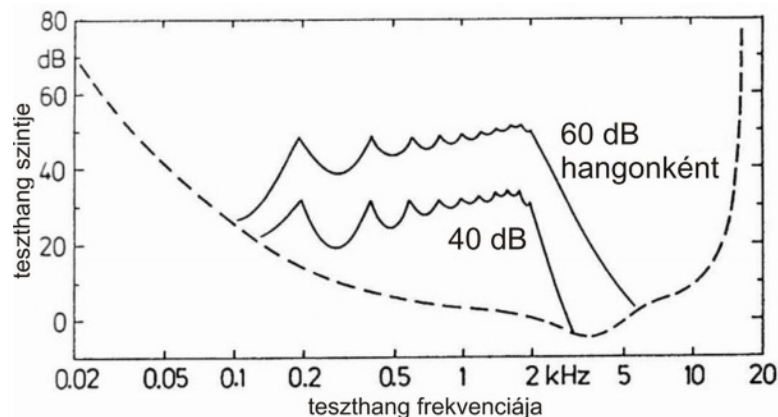
3.8. ábra. 1 kHz-es különböző szintű hangok által elfedett teszthang érzékelési szintje (folytonos vonalak) a teszthang frekvenciájának függvényében. Az abszolút hallásküszöb ebben az esetben egy vízszintes vonal 0 dB szinten. A pontozott vonalak ugyanazokat az adatokat jelenítik meg, mint a folytonos vonalak, de ezeket fordított frekvenciaskálán (felső abszcissza) jelöljük 1 kHz-re tükrözve

A 3.8. ábrán ez a jelenség látható három különböző szinten. A görbék tükrözve vannak az 1 kHz frekvenciára. Az eredeti adatok folytonos vonallal, a tükrözöttek pontozott vonallal vannak jelölve. Ha az elfedő hang szintje 20 dB, akkor az elfedési görbe laposabb az alacsonyabb frekvenciákon; 40 dB-en közel szimmetrikus; 60 dB-nél pedig tisztán látható

a különbség, a görbe a magasabb frekvenciákon lesz laposabb.

3.5. Tiszta hangok elfedése összetett hangokkal

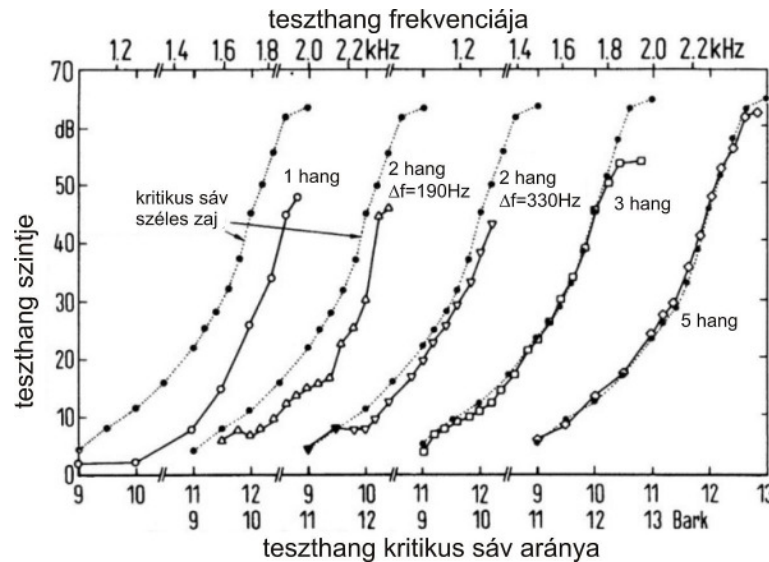
Tiszta hangokkal könnyű dolgozni, azonban a természetben ritkán fordulnak elő. Csak néhány madár éneke és a fuvola hangja tekinthető tisztának. A zenében a legtöbb hangszer hangja egy alaphangból és több felhangból tevődik össze. A különböző hangszerek által előállított eltérő hangszínek felhangjaiknak frekvenciaspektrumától függnak. A trombita például sok felhangot produkál, ezért sokkal szélesebb elfedési hatást eredményez mint a fuvola. A 3.9. ábrán tiszta hangok küszöbeit láthatjuk, amelyeket összetett hangok fednek el. Az összetett hangok 200 Hz-es alaphangból és kilenc felhangból állnak, amelyeknek azonos az amplitúdójuk, de a fázisuk véletlenszerű. A logaritmikus frekvenciaskálán a felhangok közötti távolság nagy az alacsony frekvenciákon, de a kilencedik és tizedik között már nagyon kicsi. Tehát a felhangok közötti elhajlás egyre kisebb a teszthang frekvenciájának növekedésével. A maximum és a minimum az 1.5 kHz és 2 kHz közötti frekvenciatartományban már nehezen különböztethető meg. A zenében sok összetett hangot használnak egy időben, ezek mindegyike több harmonikusból tevődik össze.



3.9. ábra. 200 Hz tíz harmonikusa által elfedett teszthang szintje a teszthang frekvenciájának függvényében. A paraméter a harmonikusok szintje

A keskenysávú zajjal és a tiszta hanggal történő elfedések különbségeket mutatnak annak ellenére, hogy ugyanazon a szinten és ugyanazzal a középfrekvenciával vizsgáljuk őket. A zajt közelíthetjük viszonylag kis számú tiszta hanggal. A 3.10. ábrán láthatunk erre a közelítésre példákat. A középfrekvencia 2 kHz, a szint 70 dB minden esetben. A kritikus sáv szélesség ezen a frekvencián 330 Hz. Először vizsgáljuk meg a közelítést egyetlen hanggal, melynek frekvenciája 2 kHz (az ábrán balra), majd két hanggal 1910 és 2100 Hz-en, majd ugyancsak két hanggal, de most 1840 és 2170 Hz-en (a két frekvencia távolsága éppen egy kritikus sáv). Végül öt hanggal is, melyeknek frekvenciája 1840, 1915, 2000, 2080 és 2170 Hz (a jobb szélén). Az abszcisszán két skála szerint is feltüntethető a teszthang szintje, fent a frekvencia, lent a kritikus sáv szerinti Bark-skála látható. Pontozott vonallal a keskenysávú zajjal való elfedés esetén, folytonos vonallal az egy vagy több hanggal

történő elfedés esetén kapott görbék vannak jelölve. Nyilvánvalóan az egyetlen hanggal történő elfedés rossz közelítést ad. Két hang esetén már közelebb jutunk a célunkhoz, főleg akkor, ha a két hang frekvenciája közötti különbség a kritikus sáv szélességgel egyenlő. Öt hang esetén azonban majdnem teljes az egyezés, az itt fellépő különbségeket akár mérési hibáknak is tekinthetjük.

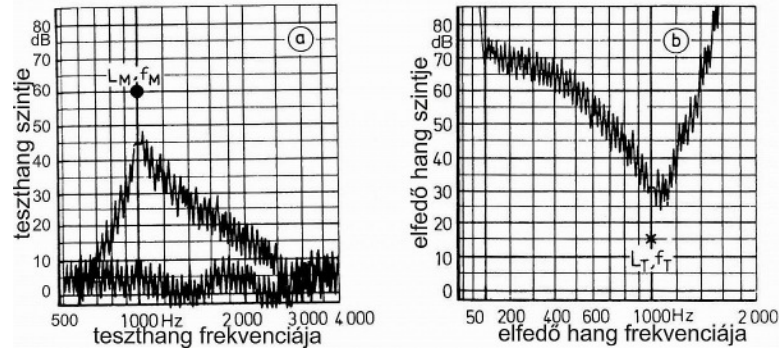


3.10. ábra. A 2 kHz-hez tartozó kritikus sávon belüli hangok által elfedett teszthang szintje a teszthang kritikus sáv arányának (alsó skála) és frekvenciájának (felső skála) függvényében. Az elfedő hangok szintje konstans 70 dB. A pontozott görbe a kritikus sáv szélességű zaj elfedése mellett mért adatokat jelöli. Ezeket az adatokat összehasonlíthatjuk a hangok alkalmazása során előállt adatokkal (folytonos vonalak). Minél több hangot használunk a keskenysávú zaj közelítésére, annál inkább egybevágnak a megfelelő görbék. Az abszcissa skáláját a görbéknek megfelelően toltuk el

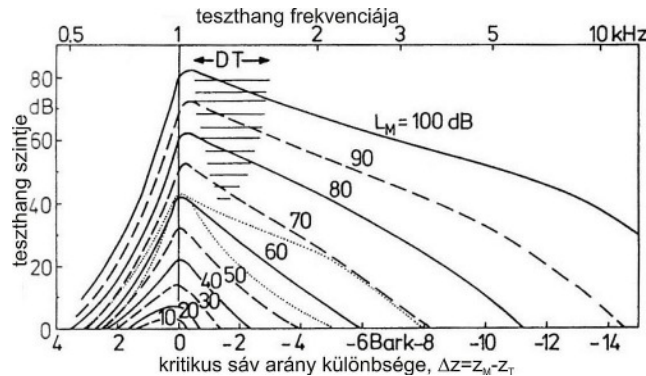
3.6. Pszichoakusztikai hangolási görbék

Amikor hangokat más hangokkal fedünk el, akkor az elfedés eredményét több módon is ábrázolhatjuk. Az elfedési jelenségnek négy paramétere van: a teszthang frekvenciája és szintje, valamint az elfedő hang frekvenciája és szintje. Általában a teszthang küszöbét ábrázolják a frekvenciájának függvényeként. Erre példa a 3.11. ábra (a) része, ahol az értékeket a követési eljárással határozták meg. A pszichoakusztikai hangolási görbék (psychoacoustical tuning curves) másfajta sémát valósítanak meg. Itt az elfedés az elfedő hang frekvenciájának függvényeként van ábrázolva. A görbét a követési eljárás segítségével is meg lehet határozni, ezt láthatjuk a 3.11. ábra (b) részén. Ebben az esetben az eljárás a következőképpen módosul: a vizsgált személy a teszthangot hallgatja, de az elfedő hang szintjét módosítja, nem a teszthangét. A klasszikus értelemben vett elfedési görbéhez képest a hangolási görbe annak inverzeként jelenik meg. A görbe jellemzője egyrészt az, hogy a meredeksége a magasabb frekvenciákon nagyobb mint az alacsonyabb frekvenciákon.

ciákon, másrészt a minimum egy kicsivel a teszthang frekvenciája fölött van (csillaggal jelölve).

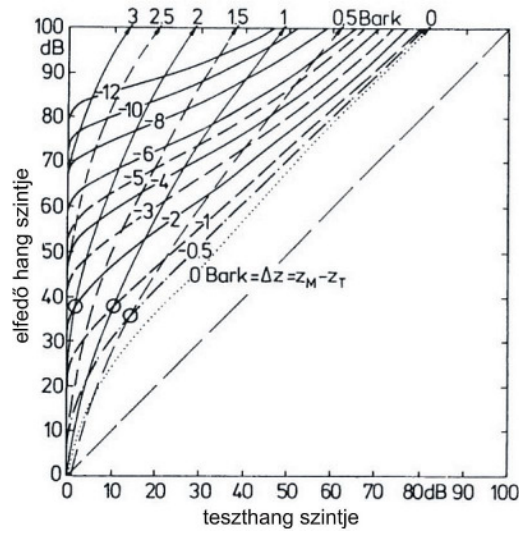


3.11. ábra. Példa a követési eljárás használatára. A követési eljárással abszolút hallásküszöböt és elfedési küszöböt mérhetünk. A görbéket ábrázolhatjuk a teszthang (a) és a pszichoakusztikai hangolási görbe (b) frekvenciájának függvényeként. Az (a) ábrán láthatjuk a teszthang frekvenciájának függvényeként ábrázolt teszthang szintet, ha azt egy elfedő hang (ponttal jelölve) fedi el. A (b) ábra ordinátája az elfedő hang szintje, az abszcissza pedig az elfedő hang frekvenciája; a teszthang szintje és frekvenciája konstans, értékeit a csillag jelöli

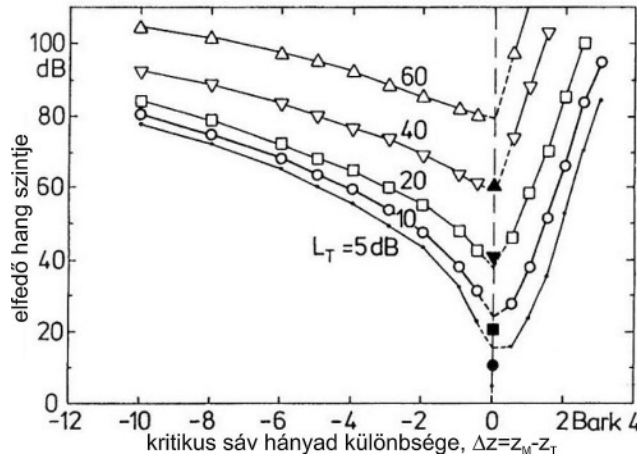


3.12. ábra. Normalizált klasszikus elfedési minták: különböző szintű elfedő hangok által elfedett teszthang szintje. Az abszcissza vagy a teszthang frekvenciája (felső skála) vagy a kritikus sáv arányok különbsége (alsó skála). A két vékony pontozott görbe a két vizsgált személy esetén mért adatok átlagtól való eltérését mutatják 60 dB elfedő szint mellett. A vonalkázott rész azt a területet jelöli, ahol olyan eltérő hangok (DT) hallhatóak, amelyek zavarják a méréseket

A szokásos normál elfedési görbék azok, amelyek egy tiszta elfedőhanggal elfedett teszthang küszöbét mutatják a teszthang frekvenciájának függvényeként és az elfedő hang szintjével mint paraméterrel. Ezek a görbék azonban függenek az elfedő hang frekvenciájától. Ez a probléma kiküszöbölhető, ha a Bark-skálát használjuk abszcisszaként a logaritmikus frekvencia skála helyett. Ekkor a teszthang frekvenciáját áttranszformáljuk kritikus sávok különbségévé, jelölje ezt Δz , amely az elfedő hang frekvencia kritikus sáv arányának (z_M -nek), és a teszthang frekvencia kritikus sáv arányának (z_T -nek) különbsége. Az így kapott adatok átlagát ábrázoltuk a 3.12. ábrán. Az áttranszformálás menete



3.13. ábra. Ugyanazokat az adatokat ábrázolja, mint a 3.12. ábra. Az ordináta az elfedő hang szintje, az abszcissza a teszthang szintje, a paraméter pedig az elfedő hang kritikus sáv arányának (z_M) és a teszthang kritikus sáv arányának (z_T) különbsége (Δz). Ilyen görbét használhatunk a 3.14. ábrához hasonló pszichoakusztikai hangolási görbék konstruálásához. A körök olyan pozíciókat jelölnek, ahol Δz pozitív és negatív értékei keresztezik egymást



3.14. ábra. Normalizált pszichoakusztikai hangolási görbék, vagyis azok az elfedő hang szintek, amelyek éppen csak elfedik az adott szintű teszthangot (paraméter). Az abszcissza a kritikus sáv arányok különbsége. A sötét szimbólumok azokat a teszthang szinteket jelölik, amelyeket a megfelelő hangolási görbe előállításához használtunk

látható a 3.13. ábrán, ahol a paraméter Δz . A paraméter pozitív és megfelelő negatív értékeivel előálló görbék metszik egymást az elfedő hang 35 dB és 40 dB körüli szintjein (ez látható az ábrán körrel jelölve). Ez azt jelenti, hogy az elfedési görbék szimmetrikusak ezeken a szinteken. Az elfedő hang 40 dB feletti értékein Δz negatív értékei olyan teszthang szintekhez vezetnek, amelyek nagyobbak, mint amelyekhez a pozitív Δz értékek vezetnek. Például ha nézzük a 60 dB-es elfedő hang szintet, az 27 dB-es teszthang szinthez vezet $\Delta z = -2$ Bark esetén, de 8 dB-es szinthez $+2$ Bark esetén. A 40 dB alatti

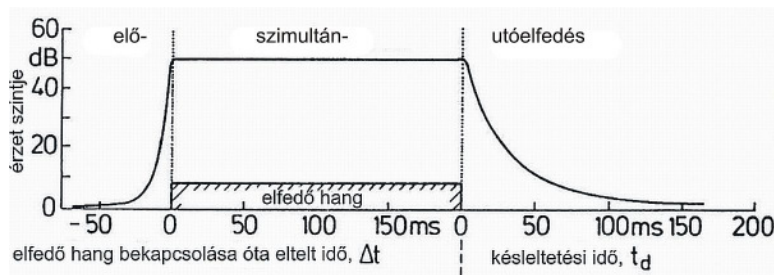
elfedő hang szinteken ez felcserélődik, ami azt eredményezi, hogy az elfedés meredeksége kisebb az alacsonyabb frekvenciákon, mint a magasabb frekvenciákon.

A pszichoakusztikai hangolási görbék akkor olvashatók le a legkönnyebben a 3.13. ábráról, ha a teszthang szintjét konstansnak hagyjuk egy viszonylag alacsony értéken. Ez az érték gyakran kevesebb, mint 20 dB. A 3.14. ábra normalizált pszichoakusztikai hangolási görbék egy halmazát mutatja, azaz az abszcissa a kritikus sáv arányok különbsége, Δz . A paraméter a teszthang szintje, L_T . A görbék aszimmetrikusak, az aszimmetria a teszthang szintjének növekedésével egyre inkább növekszik. Alacsony frekvenciákon, azaz Δz negatív értékein a görbék párhuzamosan futnak, ugyanez igaz a pozitív értékekre is. Azonban a teszthang alacsony szintjein a görbék meredeksége nagyobb, mint a magasabb szinteken.

4. fejezet

Időbeli hatások

Az előző fejezetben a teszthangokat és elfedési hangokat tartósnak és állandó állapotúnak feltételeztük. Azonban az információtovábbítás a zenében vagy a beszédben erős időbeli struktúrával rendelkezik. Az erős hangokat gyengék követik és fordítva. A beszédben az erősebb hangokat a magánhangzók reprezentálják, míg a mássalhangzók gyengébbek. A zárhang (b, d, g, gy, p, t, k, ty) tipikus példája annak, hogy egy hangot az előtte lévő hangos magánhangzó elfed.



4.1. ábra. Ez a sematikus rajz azokat a szakaszokat mutatja és jellemzi, amelyekben az előelfedés, a szimultán elfedés és az utóelfedés jelentkezik. Az utóelfedés más időskálát használ, mint az előelfedés és a szimultán elfedés

Ezt a jelenséget úgy lehet mérni, hogy veszünk korlátos időtartamú elfedési hangokat, és az elfedést rövid teszthang impulzusokkal tesztljük. A rövid jel helyzetét időben módosíthatjuk az elfedő hanghoz képest. Az így kapott eredményeket láthatjuk a 4.1. ábrán, ahol a 200 ms-os elfedő hang egy rövid hangimpulzust fed el, melynek időtartama jelentéktelen az elfedő hang időtartamához képest. Ilyen esetekben hasznos lehet két különböző időskálát használni egyszerre: az első jele legyen Δt , ez jelöli az elfedő hang kezdetéhez képest eltelt időt. Lehet negatív vagy pozitív attól függően, hogy az adott időpontban jelen van az elfedő hang, vagy sem. A másik időskála az elfedő hang elhallgatásakor indul. Ezt az időt gyakran késleltetési időnek (delay time) nevezzük, és t_d -vel jelöljük. Ordinátaként nem a teszthang hangnyomás szintjét, hanem a küszöbe feletti szintjét használjuk. Erre a szintre érzeti szintként (sensation level) hivatkozhatunk.

Az elfedésnek három különböző szakaszát különítik el attól függően, hogy az elfedő hang jelen van-e.

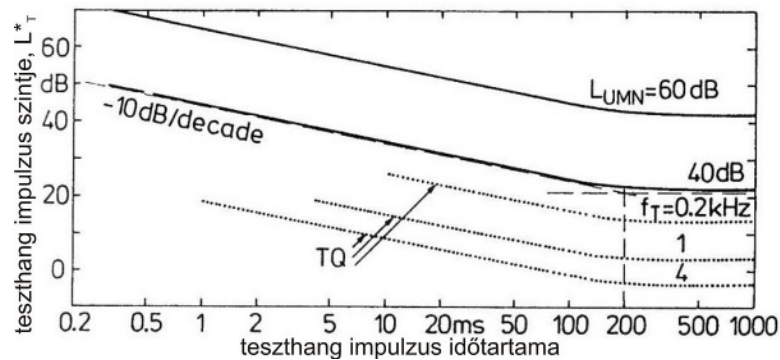
1. Előelfedésről (premasking) akkor beszélünk, amikor az elfedő hang még nincs bekapcsolva. Ebben az időintervallumban találhatók Δt negatív értékei.
2. Szimultán elfedés (simultaneous masking) az a szakasz, amikor a teszthang és az elfedő hang egyszerre hallható. Ekkor lesz Δt pozitív.
3. Miután az elfedő hang véget ér, következik az utóelfedés (postmasking). Itt találhatók t_d pozitív értékei, amikor az elfedő hang nincs jelen fizikailag, mégis hatást gyakorol a teszthangra.

Az utóelfedés nem meglepő jelenség, hiszen az elfedő hang nem hallgat el egyik pillanatról a másikra, bizonyos lecsengési időre van szüksége. Az előelfedés viszont már meglepő jelenség, mivel még az elfedési hang bekapcsolása előtt jelentkezik. A magyarázat erre az, hogy ugyanúgy, mint ahogy nem hallgat el a hang egyik pillanatról a másikra, úgy nem is jelenik meg, tehát szükség van egy kis időre, amíg felépül. A hangos elfedési hangok esetén gyors, a gyengébbek esetén lassú felépülési időre van szükség. Az előelfedés időtartama viszonylag rövid, csak körülbelül 20 ms. Az utóelfedés azonban több mint 100 ms-nál is tovább tarthat. Így aztán az utóelfedést tekinthetjük a domináns nem-szimultán elfedési hatásnak.

4.1. Szimultán elfedés

Mind az abszolút hallásküszöb, mind az elfedési küszöbök függenek a teszthang időtartamától. A mérések elvégzéséhez nagyon rövid teszthangokkal kell dolgozni, ezért ezeket a függőségeket figyelembe kell venni, hogy időbeli hatásokról beszélhessünk nem-szimultán elfedés esetén. Kétféle függőséget lehet megkülönböztetni: az egyik a küszöbök függése egy teszthang időtartamától, a másik az ismétlődő rövid teszthangok ismétlési arányától való függés. A 4.2. ábra az első függőséget szemlélteti szinuszos teszthang impulzusok esetén. A pontozott vonalak a különböző frekvenciákhoz tartozó hallásküszöböt jelölik, a folyamatos vonalak pedig a hangimpulzusok küszöbét, amelyeket 40 dB-es és 60 dB-es egyenletes elfedési zajjal fedtünk el. Később majd láthatjuk, hogy ezek nem függenek a frekvenciától. Mindkét függőség azonos a hallásküszöb és az elfedési küszöb esetén. Az időtartamtól való függés azt mutatja, hogy a teszthang küszöbe konstans 200 ms-nál nagyobb időtartamok mellett. A 200 ms-nál rövidebb időtartamok esetén a hallásküszöb és az elfedési küszöb az időtartam növekedésével arányosan csökken tízszeres időtartam alatt 10 dB-el. A hatás frekvenciafüggőségét a különböző hallásküszöb-görbék jelzik. Egyenletes elfedési zaj mellett az elfedési küszöb független lesz a frekvenciától (lásd 3.2. ábra), a két folytonos vonal tehát a teljes frekvenciatartományra vonatkozik.

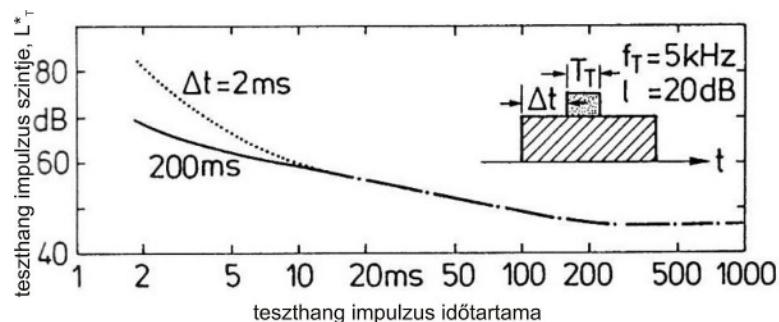
Az ismétlési aránytól való függőség esetén hasonló megállapításokat tehetünk. A 200 ms-os intervallum 5 Hz-es ismétlési aránynak felel meg. Ezért azt mondhatjuk, hogy a küszöbök függetlenek az ismétlési aránytól, ha az kisebb, mint 5 Hz, de csökkennek nagyobb ismétlési arányok esetén, míg végül el nem érjük azt, hogy a teszthang állandóan szó, azaz az ismétlési arány olyan magas, hogy az impulzusok között nincs szünet. Ez 5 ms-os hangimpulzusok esetén 200 Hz-es ismétlési arány mellett következik be.



4.2. ábra. Adott szintű egyenletes elfedési zajok elfedése mellett éppen hallható teszthang impulzus szintje (L_T^*) az impulzus időtartamának függvényében. Az abszolút hallásküszöböt (TQ) pontozott görbék jelölik három különböző frekvencia esetén. Megjegyezzük, hogy L_T^* annak a folytonos hangnak a szintje, amelyből a teszthang impulzust származtatjuk. A vízszintes szaggatott vonal aszimptotát jelöl

A teszthang időtartamának csökkentése frekvenciatartományának bővítését eredményezi. Ez a hatás a legrövidebb használható időtartamot egy olyan értékben határozza meg, amely mellett a frekvenciatartomány megfelel a kritikus sáv szélességnek. Az időbeli hatások mérésére általában haranggörbe alakú hangimpulzusokat használnak.

A hallásküszöbre vonatkozó mérések során ténylegesen csak egy kis frekvenciatartományt vizsgálunk, egészen pontosan ott, ahol a hallásunk a legérzékenyebb. Ez azt jelenti, hogy a szélessávú zajok esetén mért küszöböt valójában egy keskenysávú teszthang segítségével határozzuk meg, 3 kHz körüli frekvenciatartományban. Ezért a hallásküszöb szélessávú zaj mellett hasonló egy keskenysávú zaj esetén mért hallásküszöbhez, ahol a keskenysávú zaj középfrekvenciája hallásunk legérzékenyebb frekvenciatartományának közelében van. A zaj azon komponenseit, melyek a legérzékenyebb frekvenciatartományon kívül esnek, nem vesszük figyelembe.

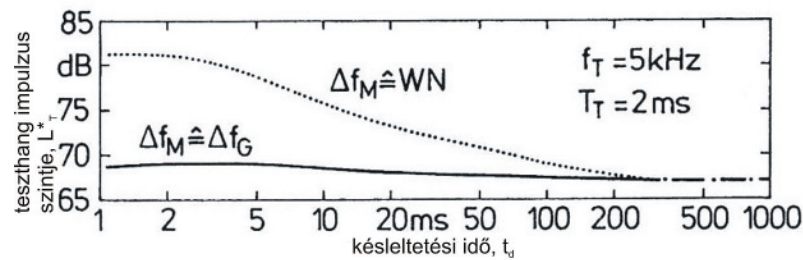


4.3. ábra. 2 ms-os (pontozott) vagy 200 ms-os (folytonos) teszthang impulzusok szintje egy 20 dB-es fehér zaj bekapcsolása után. A görbék az 5 kHz-es teszthang időtartamának függvényeként vannak ábrázolva. Felhívjuk a figyelmet a rövid teszthang impulzusok esetén jelentkező szintkülönbségre (túlfutási hatás)

Az ismétlési aránytól való függés hasonló hatásokat mutat mint a hangok esetében

tárgyaltak. Az elfedési küszöb konstans marad a nagyon kicsi ismétlődési arányokra, egészen 10 Hz-ig. Nagyobb ismétlési arányok mellett ismét csökken, majd végül elér egy nyugalmi állapotot.

A szimultán elfedés tartományán belül további hatások jelentkeznek röviddel az elfedő hang bekapcsolása után. A 4.3. ábrában lévő rajzon egy ilyen időbeli struktúra látható. Egy magas frekvenciájú (5 kHz) teszthangot használunk rövid teszthang impulzusok előállítására. Az elfedő hangimpulzus fehér zajból származik, és a teszthang impulzus időtartamánál hosszabb ideig tart. Az elfedő hang 20 dB-es sűrűségi szintje mellett előálló elfedési küszöb a 4.3. ábrán látható, ahol az elfedési küszöb a teszthang impulzus időtartamának függvényeként van ábrázolva. Ha a teszthang impulzus 200 ms-al az elfedő hang kezdete után jelentkezik, akkor nincs különbség az előzőekben tárgyalt elfedési küszöbhez képest. Azonban ha ez a Δt érték nagyon kicsi, azaz például 2 ms, akkor az elfedési küszöb a rövid időtartamok esetén nagy eltérést mutat: a küszöb 10 dB-el magasabb 2 ms-nál. Ez a jelenség csak a körülbelül 10 ms-nál rövidebb időtartamok esetén jelentkezik.



4.4. ábra. A túlfutási hatás, vagyis egy 5 kHz-es 2 ms időtartamú teszthang impulzus szintje a késleltetési idő függvényeként. Az elfedő 25 dB szintű és 5 kHz frekvenciájú kritikus sáv szélességű zaj (folytonos görbe) vagy fehér zaj (pontosított görbe)

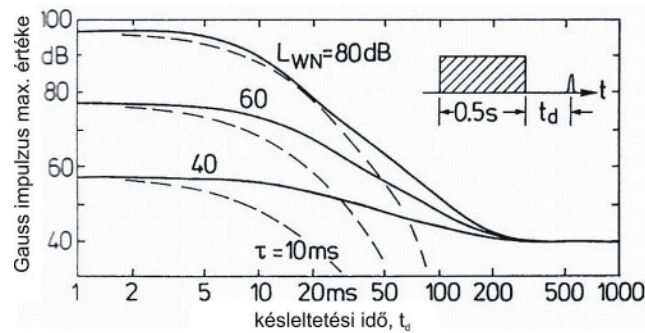
Ez az úgynevezett túlfutási hatás (overshoot effect) függ még az elfedő hang és a teszthang spektrális összetételétől is. A jelenség megszűnik, ha a két hangnak hasonló a spektruma. A 4.4. ábra az elfedő hang két különböző sáv szélességére mutat példát. Az első esetben az elfedő hang fehér zaj (mint az előző ábrán), ennek eredménye pontosított vonallal van jelölve. A második esetben a folyamatos vonal mutatja egy 5 kHz középfrekvenciájú kritikus sáv szélességű elfedő hang által előállított eredményeket. A késleltetési idő, Δt , mely az elfedő hang kezdete és a 2 ms-os teszthang impulzus között telik el, az abszcissza. A túlfutási hatás szélessávú elfedő hangok esetén nagyon kis Δt értékeken a legnagyobb. Nagyobb Δt értékek esetén csökken, és teljesen eltűnik körülbelül 200 ms után. A hatás keskenysávú elfedő hangok esetén egyáltalán nem jelentkezik.

4.2. Előelfedés

Azt a szakaszt, ahol az előelfedés jelentkezik, láthattuk a 4.1. ábrán. Az előelfedést főként különálló elfedő hangimpulzusokkal mérhetjük, ezek az eredmények azonban bizonytalanok. Azt még nem sikerült eldönteni, hogy vajon az előelfedés függ-e az elfedő hang időtartamától. Az előelfedés körülbelül 20 ms-ig tart minden esetben. Ez azt jelenti,

hogy a küszöb változatlan marad, míg el nem éri a $\Delta t = -20$ értéket. Ezután a küszöb emelkedik, majd eléri a szimultán elfedésnél található szintet, körülbelül akkor, amikor az elfedő hangot bekapcsoljuk. Az előelfedésnek nincs jelentősége, mivel csak 20 ms-ig tart, ezért általában nem vesszük figyelembe.

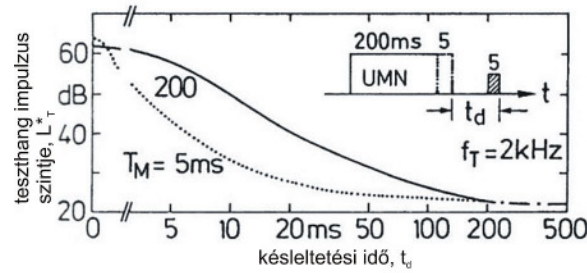
4.3. Utóelfedés



4.5. ábra. Utóelfedés: egy éppen hallható 20 ms-os Gauss impulzus maximális szintje a készletelési idő függvényeként (abszciissa) az adott szintű fehér zaj elfedő kikapcsolása után. A szaggatott vonalak az exponenciális lecsengések görbéi

A fehér zaj elfedő hanggal előállított utóelfedés spektrális hatások nélkül mérhető, ha egy rövid Gauss-impulzust használunk teszthangként. A 4.5. ábra ordinátáját a Gauss-impulzus maximum értéke adja szintértékben kifejezve. Ez az ábra azt mutatja, mennyi a szükséges szint ahhoz, hogy éppen elérjük a küszöböt az utóelfedésben. A küszöb az elfedő hang végétől számított készletelési idő, t_d függvénye. A paraméter a fehér zaj elfedő hang teljes szintje. A folytonos vonalak azt mutatják, hogy szinte semmi csökkenés nem jelentkezik az elfedő hang kikapcsolása utáni első 5 ms-ban. Az értékek megegyeznek a szimultán elfedés során mért értékekkel. Körülbelül 5 ms készletelési idő (t_d) után a küszöb csökkenni kezd, és körülbelül 200 ms-nál eléri az abszolút hallásküszöböt. A szaggatott vonalak 10 ms-os idejű exponenciális lecsengéseket mutatnak. A szaggatott vonalak és a folytonos vonalak összehasonlítva megállapíthatjuk, hogy az utóelfedés hatása nem írható le exponenciális lecsengésként.

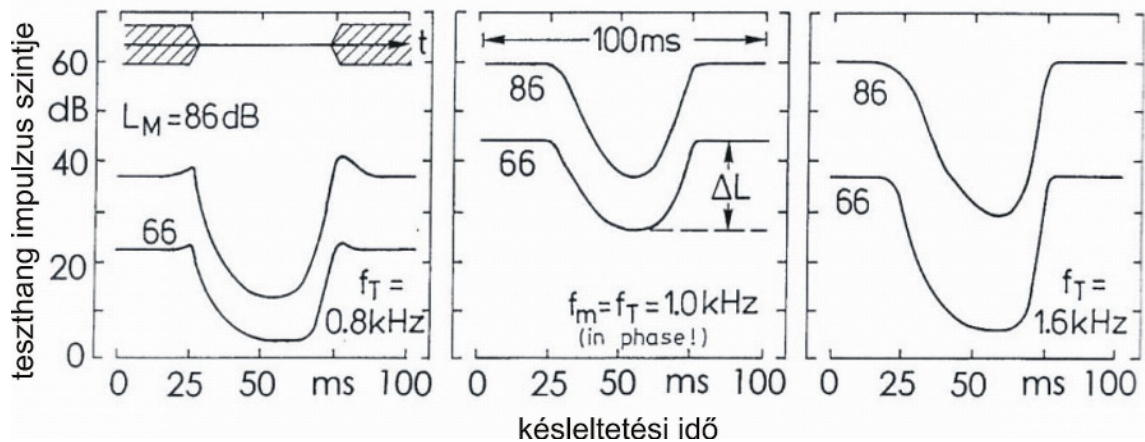
Az utóelfedés függ az elfedő hang időtartamától. Ez kiderül, ha a teszthang egy 2 kHz-es hangimpulzus 5 ms-os időtartammal, ahogy az a 4.6. ábrán is látható. Az abszciissa ismét az elfedő hang vége után eltelt idő. Az ordináta a teszthang impulzus szintje. Egy 200 ms időtartamú elfedő hang esetén a folytonos görbe által jelölt utóelfedés hasonló a 4.5. ábrán lévő utóelfedéshez. A pontozott vonal egy olyan utóelfedést ábrázol, ahol az elfedő hang csak 5 ms-ig tart. Ekkor a csökkenés kezdetben sokkal meredekebb. Ez azt mutatja, hogy az utóelfedés nagymértékben függ az elfedő hang időtartamától, ezért ez egy erősen nemlineáris jelenség.



4.6. ábra. Az utóelfedés függ az elfedő hang időtartamától: éppen hallható teszthang impulzusok szintje késleltetési idejüknek függvényeként a 200 ms-os és az 5 ms-os elfedő hang kikapcsolása után. Az egyenletes elfedési zaj szintje 60 dB; a 2 kHz-es teszthang időtartama 5 ms. A késleltetési idő ebben az esetben az elfedő hang vége és a teszthang vége között eltelt idő

4.4. Időbeli elfedési minták

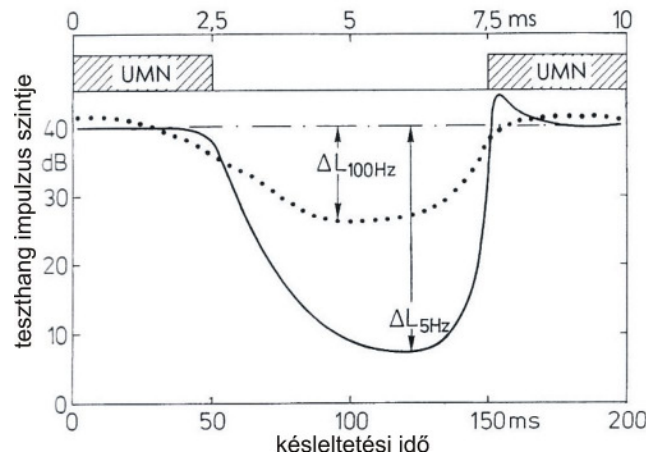
Időbeli elfedési minták (temporal masking patterns) előállítása olyan elfedő hangokkal történik, amelyeknek időbeli burkolójuk jellegzetes struktúrával rendelkezik. Az amplitúdó modulált hangok vagy zajok, frekvencia modulált hangok vagy a keskenysávú zaj burkolójának ingadozása tipikus példái azoknak a hangjelenségeknek, amelyek időbeli elfedési mintákat állítanak elő.



4.7. ábra. Egy 10 Hz-es négyzögjellel amplitúdó modulált 1 kHz-es hang által előállított időbeli elfedési minták. Adott frekvenciájú 3 ms-os teszthang impulzusok érzékelési szintjét ábrázoltuk a késleltetési idő függvényeként. Az abszcissza a 10 Hz-nek megfelelő 100 ms-os elfedő periódus felosztása. A paraméter az 1 kHz-es elfedő hang szintje; mindegyik panel különböző teszthang frekvenciának (f_T) felel meg

Az időbeli elfedési minták három csoportját a szimultán elfedés, az előelfedés és az utóelfedés képi. Mindhárom elfedési jelenség függ az elfedő hang spektrumától. Egy szélessávú zaj elfedő hang esetén az elfedési jelenség frekvenciaválasztása másodlagos szerepet játszik, keskenysávú zaj esetén azonban már nagyon fontos. Az elfedési minta alacsonyabb és magasabb frekvenciákon lévő meredekségein az elfedési küszöb sokkal a-

lacsonyabb, mint a keskenysávú zaj elfedő hang frekvenciatartományában. Az időbeli minta alakja hasonló minden frekvencián. A 4.7. ábra egy 1 kHz-es elfedési hang által előállított időbeli elfedési mintát mutat, ahol az elfedési hangot egy 2 ms-os emelkedéslengésű haranggörbe alakú hullámmal moduláljuk. A moduláció frekvenciája eléri a 10 Hz-et, tehát a 100 ms-os periódus fel van osztva egy 50 ms-os részre, ahol az elfedő hang hallható, és egy 50 ms-os szünetre. Az ábrán az időskálát úgy választottuk, hogy a szünet minden diagramnak a közepén helyezkedjen el. Az ordináta az adott frekvenciájú (f_T) 3 ms-os teszthang impulzus szintje. A teszthang frekvenciája a három ábrán különböző, a paraméter minden esetben az elfedő hang szintje. A 0.8 kHz-es és 1 kHz-es teszthang frekvenciák esetén az időbeli elfedési mintákat jelző folyamatos vonalak nagyon hasonlóak, körülbelül a két elfedő hang közötti szintkülönbséggel vannak eltolva egymáshoz képest. Azonban az 1.6 kHz frekvenciájú teszthang esetében a távolság a két időbeli elfedési minta között - különösen a szünet alatt - nagyobb. Ez abból következik, hogy az elfedés során a magasabb frekvenciájú meredekség nemlineárisan függ a szinttől, ahogy erről a 3.4. szakaszban már volt szó. Minden időbeli elfedési minta jelzi az előelfedés viszonylag meredek emelkedését, és az utóelfedés viszonylag lassú csökkenését. Azonban az utóelfedés csökkenése sokkal rövidebb, mint ahogy az a 4.5. ábra alapján várható lenne. Ez a különbség a 4.6. ábra által ábrázolt jelenségből érthetővé válik; az utóelfedés függ az elfedő hang időtartamától. A 4.7. ábrán ábrázolt időbeli elfedési mintáknál az elfedő hang időtartama csak 50 ms, azaz sokkal rövidebb, mint a 4.6. ábrán szereplő 200 ms-os elfedő hang, és ezért meredekebb az utóelfedés csökkenése.



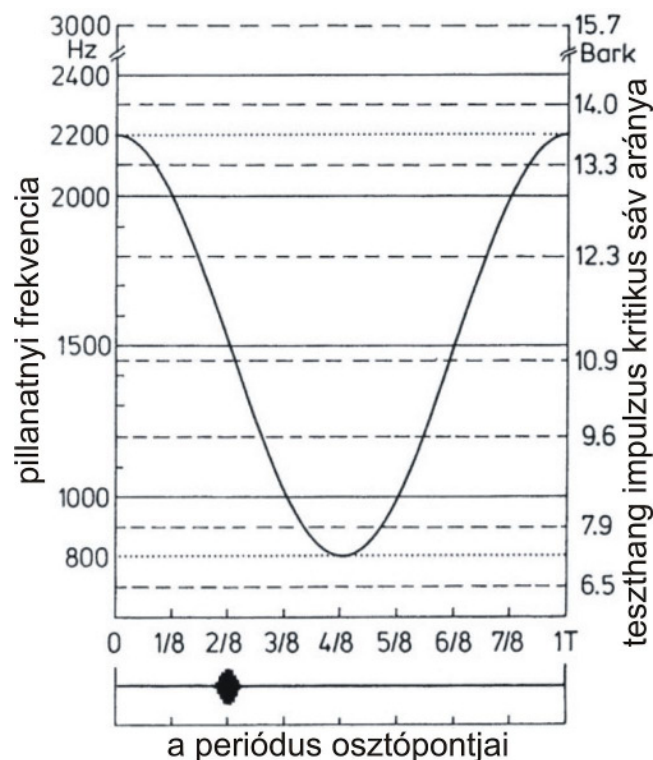
4.8. ábra. Egyenletes elfedési zaj által előállított időbeli elfedési minták 5 vagy 100 Hz-es modulációs frekvenciájú négyzetjellel modulálva. A teszthang impulzusok 3 ms-osak a teszthang frekvenciája 3 kHz

A 4.8. ábra egy egyenletes elfedési zaj időbeli elfedési mintáit mutatja, amely zajt moduláltuk 5 Hz-es (folyamatos vonal, alsó abszcissza), valamint 100 Hz-es (pontosított vonal, felső abszcissza) négyzetjellel.

Alacsony modulációs frekvencia esetén (5 Hz) a különbség az időbeli elfedési minta maximuma és minimuma között (ΔL) meghaladja a 30 dB-t is. A magasabb modulációs frekvencia esetén (100 Hz) a szintkülönbség (ΔL) jelentősen csökken, főként az utóelfedés

hatása miatt. Megállapítható tehát, hogy az elfedési hangban lévő rövid kihagyások „áthidalhatóak” az utóelfedés hatásával.

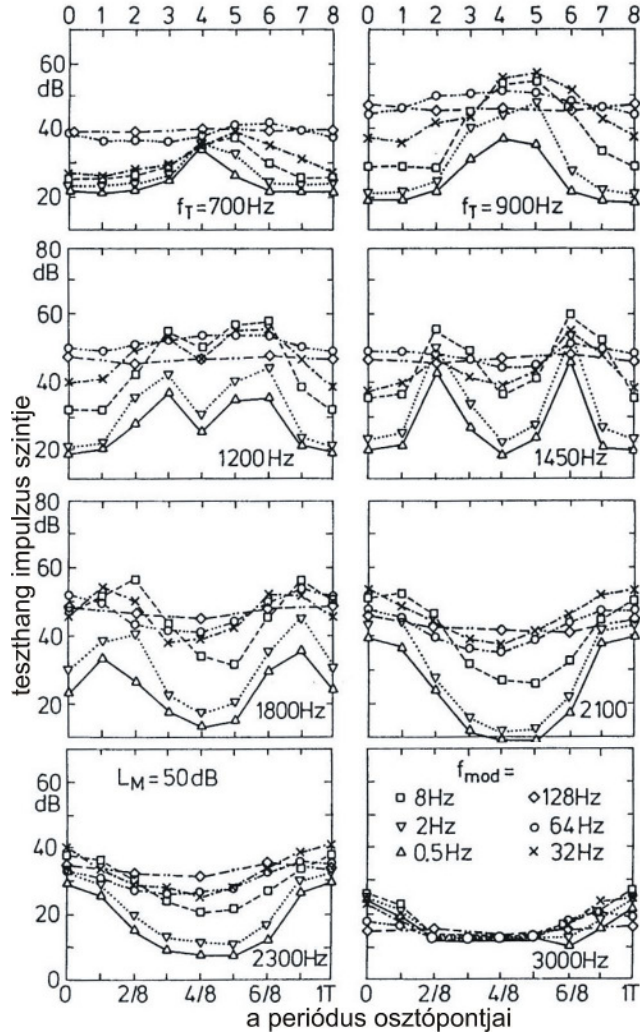
A frekvenciában periodikusan modulált hangok az amplitúdómodulált hangokhoz hasonló módon állítanak elő időbeli elfedési mintákat. Erre példa a szinuszosan frekvenciamodulált tiszta elfedő hang által előállított időbeli elfedési minta. Ekkor a küszöböt olyan rövid hangimpulzusokkal kell mérnünk, amelyek a frekvenciamoduláció adott pillanatában jelentkeznek. A teszthangimpulzusok frekvenciáit úgy válasszuk meg, hogy azok legalább a frekvenciaingadozás teljes tartományát lefedjék. A paraméterek a teszthang impulzus időbeli pozíciója a moduláció periódusán belül, a teszthang impulzus frekvenciája, a moduláció frekvenciája és az elfedő hang szintje.



4.9. ábra. Egy szinuszosan frekvenciamodulált ($\pm 700\text{ Hz}$) 1.5 kHz -es hang pillanatnyi frekvenciája az elfedő hang periódusának nyolcad egységében kifejezett idő függvényeként. A szaggatott vízszintes vonalak a teszthang impulzus frekvenciáit (bal skála) és kritikus sáv arányait jelölik (jobb skála). Az alsó részen egy 1200 Hz -es 5 ms hosszúságú teszthang impulzust jelöltünk a $2T/8$ pozícióban

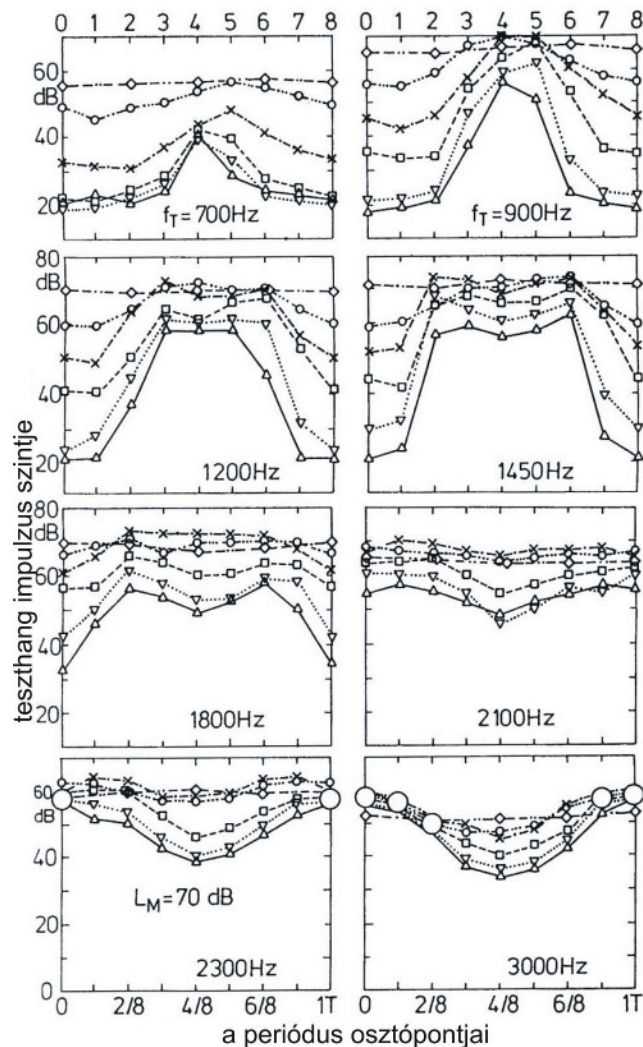
Nézzünk meg egy példát, ahol az elfedő hang egy 1.5 kHz -es szinuszosan frekvenciamodulált hang ($\pm 700\text{ Hz}$). A 4.9. ábrán láthatjuk a frekvenciamodulált hang pillanatnyi frekvenciái közötti különbséget, a moduláció periódusán belüli felosztását nyolc részre és a teszthang impulzus frekvenciáit (vízszintes szaggatott vonalak). A frekvenciaskála mellett a Bark-skála is fel van tüntetve. A moduláció periódusa a legmagasabb pillanatnyi frekvencián (2200 Hz) indul, a periódus közepén eléri a minimumát, és áthalad az 1500 Hz -es középfrekvencián, a periódus $2/8$ és $6/8$ pozícióin. Alul látható a periódus $2/8$

pozícióján egy 1200 Hz-es tesztang impulzus időfüggvénye 8 Hz-es modulációs frekvencia esetén (125 ms-os periódus).



4.10. ábra. Időbeli elfedési minták, azaz az éppen hallható tesztang impulzus szintje az intervallumon belüli temporális pozíciójának függvényeként, az elfedő periódusának (FM hang, 1.5 kHz $\pm$ 700 Hz) nyolcadaiban adva. A paraméter a 4.10. ábra jobb alsó paneljében adott szimbólum által mutatott modulációs frekvencia. Az elfedő hang szintje 50 dB a 4.10. ábrán és 70 dB a 4.11. ábrán

A 4.10. ábra egy 50 dB-es 1500 Hz-es frekvenciamodulált hang által előállított időbeli elfedési mintákat mutatja nyolc különböző tesztang frekvencia esetén. Mind a nyolc esetben a modulációs frekvencia a paraméter. Az adatok elárulják, hogy a fül követni tudja a frekvencia modulált hang által előállított elfedési hatás időbeni folyamatát körülbelül 8 Hz-es frekvencia modulációig. Ennél magasabb modulációs frekvenciák esetén a minta elveszti megszokott alakját. A 128 Hz-es modulációs frekvencia esetében a minta például majdnem teljesen lapos. Ugyanez igaz a frekvenciamodulált elfedő hang magasabb hangnyomás szintjeire is, ami a 4.11. ábrán látható. A 700 Hz-es tesztang frekvencia esetén található minták mind az 50 dB-es, mind a 70 dB-es elfedő hangoknál csak egyetlen

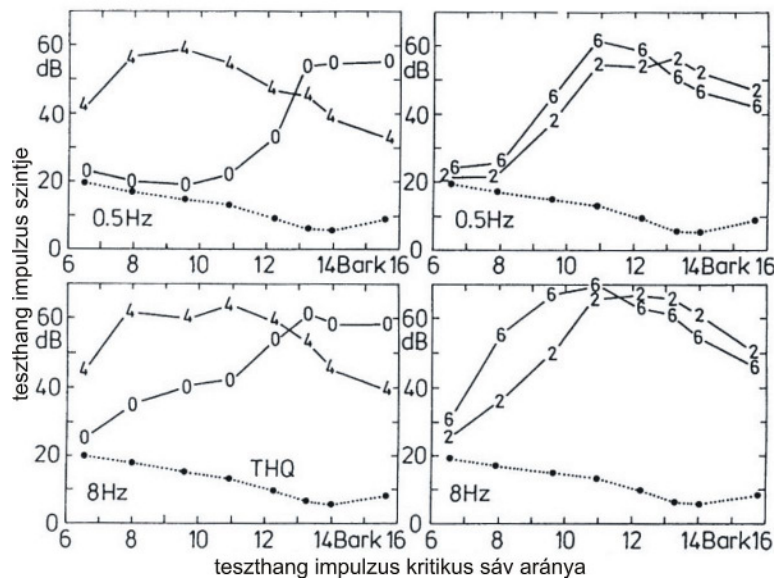


4.11. ábra. A leírás a 4.10. ábra mellett látható

csúccsal rendelkeznek a perióduson belül. A csúcst az elfedő hang legalacsonyabb pillanatnyi frekvenciája alatti elfedés meredeksége okozza. A 900 Hz-es frekvencián az időbeli elfedési minta csúcsa valamivel szélesebb, mert a modulált elfedő hang frekvenciája ezt a teszthang frekvenciát kétszer éri el rövid időintervallumon belül (lásd a 4.10. ábrát). Az 1200 Hz-es és 1450 Hz-es teszthang frekvenciákon a minták két csúcst mutatnak az alacsonyabb modulációs frekvenciákon, ami azt jelenti, hogy az elfedő hang pillanatnyi frekvenciája a teszthang impulzus frekvenciáját kétszer haladja meg a moduláció periódusán belül.

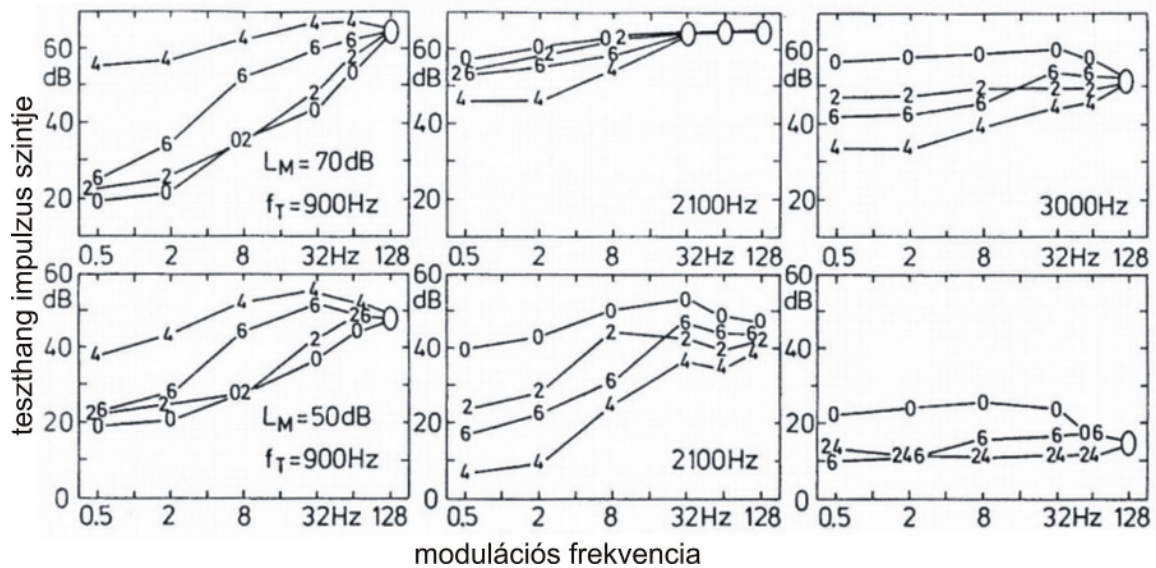
Néhány alkalmazás esetén érdekes lehet ezeket az adatokat a teszthang impulzusok frekvenciájának vagy kritikus sáv arányának függvényeként ábrázolni. A 70 dB-es frekvenciamodulált elfedő hang által előállított elfedési minták a 4.12. ábrán láthatóak két modulációs frekvencia esetén, 0,5 Hz-nél (felső ábrák) és 8 Hz-nél (alsó ábrák). A paraméter a teszthang impulzus pozíciója az elfedő hang periódusán belül. A két bal oldali ábra mutatja az elfedő hang szélsőséges pozícióinak előfordulásakor előálló elfedési küszöböt.

Ezek a pozíciók a 2000 Hz („0”-val jelölve) és a 800 Hz („4”-gyel jelölve). A két jobb oldali panel a frekvenciamoduláció zérusátmenteténél lévő adatokat mutatja, azaz az 1500 Hz-es pontot, amit a „2”-es és a „6”-os időbeli pozíciókon ér el. A minták a szélsőértékekénél („0” és „4”) a várt értékeket adják. Ugyanez igaz a zérusátmenet értékeire („2” és „6”) 0.5 Hz-es modulációs frekvencia esetén. Közelebbről vizsgálva azonban látható, hogy az utóelfedés már hat ezekre az adatokra, így a „2”-es görbe megmarad a „6”-os görbe alatt az alacsonyabb kritikus sáv arányok mellett és fordítva helyezkednek el magasabb kritikus sáv arányok esetén. A „2”-es és „6”-os görbének azonosnak kellene lennie, ha az előelfedés és utóelfedés hasonlóan viselkedett. Mivel az utóelfedés sokkal tovább tart, mint az előelfedés, ezeket a különbségeket várjuk. A 8 Hz-es modulációs frekvencia mellett ez a hatás még hangsúlyozottabbá válik, ami majdnem 20 dB-es különbségekhez vezet az elfedési küszöbön a „2”-es és „6”-os által jelzett pozíciókban az alacsonyabb kritikus sáv arányok esetén. A különbség nem ilyen nagy a magasabb kritikus sáv arányok esetén, mert az elfedési minták meredeksége magasabb frekvenciákon kisebb.



4.12. ábra. Éppen hallható teszhang impulzus szintje kritikus sáv arányának függvényében. A paraméter a teszhang impulzus időbeli pozíciója (az elfedő periódusának nyolcadaikét adottak). Az 1.5 kHz-es, ± 700 Hz-es FM elfedő modulációs frekvenciája 0.5 Hz a felső, 8 Hz az alsó két panelen. Az abszolút hallásküszöböt (THQ) pontokkal jelöltük

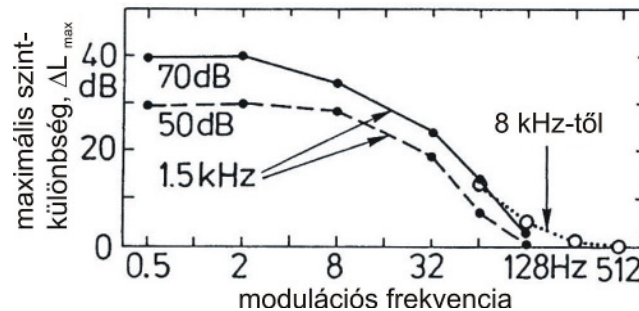
Sok időbeli hatásoktól függő érzet esetén az elfedési hang periódusán belül az elfedésben lévő eltérés fontos szerepet játszik. A 4.13. ábra elfedési küszöböket mutat a frekvenciaingadozás szélsőértékeinek pozícióinál, valamint a zérusátmeneteknél. A felső sor esetén az elfedési szint 70 dB, az alsó sor esetén 50 dB. A három frekvenciaérték a 900 Hz, egy érték valamivel a középfrekvencia felett (2100 Hz) és 3000 Hz, ami a felső szélsőérték (2200 Hz) felett van. A legalacsonyabb és a legmagasabb görbe közötti különbség minden esetben azt a szintkülönbséget jelzi, ami az elfedés után bekövetkezik az elfedő hang minden periódusán belül. Ez a különbség nyilván csökken magas modulációs frekvenciák esetén. A 128 Hz-es modulációs frekvencia körülbelül 8 ms-os periódusnak felel meg. Az



4.13. ábra. Éppen hallható teszthang impulzus szint az 1.5 kHz-es, ± 700 Hz-es FM elfedő hang modulációs frekvenciájának függvényeként. A paraméter a teszthang impulzus időbeli pozíciója az elfedő periódusának nyolcadaiban megadva. A teszthang impulzus frekvenciája minden panelen különböző. Az elfedő szintje 70 dB a felső, 50 dB az alsó sorban

5 ms-os teszthang impulzus ekkor már nagyobb, mint a periódus fele. Ez azt jelenti, hogy ezen a magas modulációs frekvencián a küszöb különbségeket valószínűleg csökkentette ez az ide nem illő reláció. További méréseket végzünk egy 8 kHz középfrekvenciájú elfedő hangon, ± 2 kHz ingadozással. Ezzel a feltétellel a csupán 1 ms-os teszthang impulzusok még mindig rendelkeznek spektrális szélességgel a kritikus sáv szélességein belül. Ezért a méréseket csak 512 Hz modulációs frekvenciáig lehet elvégezni. Ezek az adatok azt mutatják, hogy a 250 Hz-es modulációs frekvencia nagyon kis különbséget eredményez a maximális és minimális elfedés között a perióduson belül, a különbség legfeljebb 1 dB. Ezeket az adatokat felhasználva a 4.10. és a 4.11. ábrán látottakkal együtt, a 4.14. ábrát kapjuk, ahol egy perióduson belüli küszöbök közötti maximális különbség, ΔL_{max} látható 70 és 50 dB-es elfedő hangok esetén. Az ábrából kiderül, hogy a hallásunk jól tudja követni a modulációt majdnem 8 Hz-es modulációs frekvenciáig. Magasabb modulációs frekvenciák mellett a ΔL_{max} érték csökken, és 250 Hz közelében éri el a 0-t.

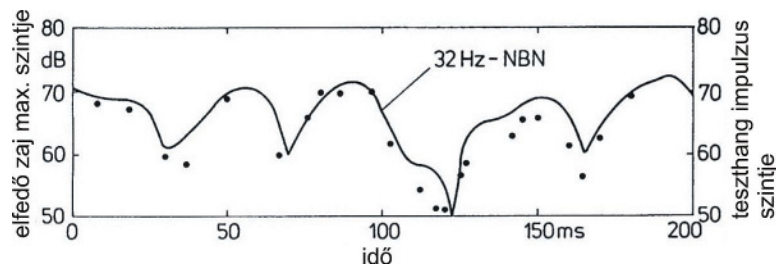
Sok hang, amely elfedést eredményez, nem rendelkezik periódussal. Ez különösen igaz a zajokra. Hogy az ilyen zajok által produkált elfedést mérhessük, egy olyan speciális zajt kell előállítani, amelynek hosszú periódusa van. Ilyen zajokat véletlenszerűen ismételt impulzusok generálnak. A zajok periódusa nem hallható, ha az hosszabb néhány másodpercnél. Egy teszthang impulzust bizonyos időpillanatban ezen mesterséges zaj periódusán belül megszólaltatva az időbeli elfedési mintáknál előállt elfedési jelenséghez hasonló jelenséget kapunk. A 4.15. ábrán egy 2 ms-os 3 kHz-es teszthang impulzus küszöbe látható egy ilyen elfedő zaj 200 ms-os intervallumában az idő függvényében. Ebben az esetben az elfedő hang egy 4 kHz-es ismétlődő zaj 32 Hz-es sáv szélességgel és 2 s-os periódussal. Az ábrán a folytonos vonal a hangnyomás burkolójának logaritmusa mutatja az idő függvényeként. A baloldali ordinátán jelölt érték a keskenysávú elfedő zaj maximális



4.14. ábra. Az 1.5 kHz-es, ± 700 Hz-es FM elfedő periódusában található éppen hallható teszthang szint maximális különbsége 50 dB és 70 dB elfedő szint esetén. A pontozott görbe egy 8 kHz-es, ± 2 kHz-es FM elfedő által előállított adatokból származik

szintje. Az elfedő zaj 32 Hz-es sávszélessége miatt az elfedő burkolójában található maximum és minimum tisztán látható. A maximum és minimum közötti időbeli átlagtávolság körülbelül 50 ms.

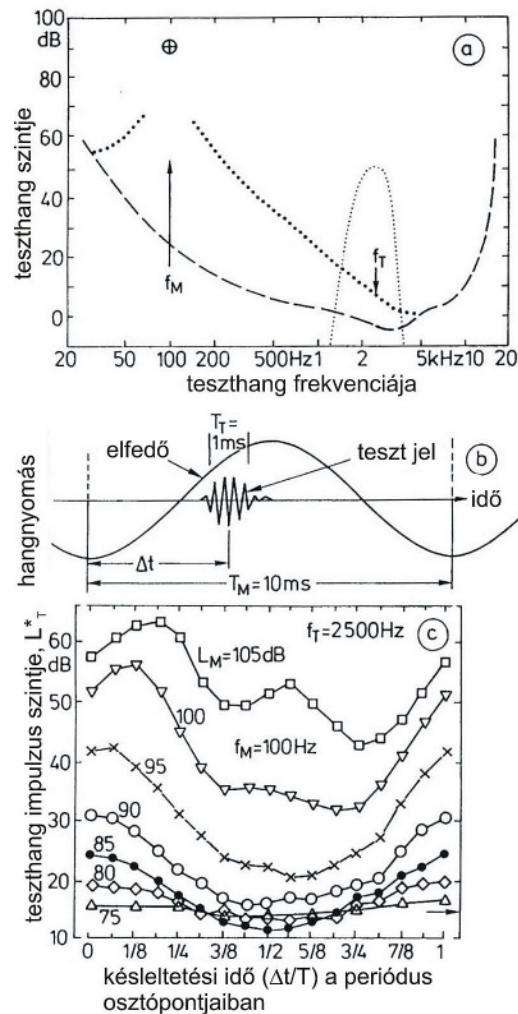
Az elfedési küszöbök az alacsony frekvenciákon 3 kHz-es hangimpulzusokkal mértük. Ezen feltétel mellett az elfedő zaj burkolójának maximum szintjei majdnem megegyeznek a teszthang impulzus azon hangnyomás szintjeivel, amelyek ahhoz szükségesek, hogy a hang éppen hallható legyen (az ábrán ezeket a pontok jelzik, a mennyiségek a jobboldali ordinátán láthatóak). Összehasonlítva a folytonos vonalat és a pontokat, amelyek az elfedési küszöbököt reprezentálják, egyezést láthatunk a két érték között, és nem csak az átlagban, hanem az elfedő zaj burkolójának időbeli struktúrájában is. Szélesebb zajsávok esetén a burkoló gyorsabban változik, ezért az utóelfedés a legtöbb völgyet kitölti.



4.15. ábra. Egy 4 kHz-es középfrekvenciájú 32 Hz szélességű zaj által előállított időbeli elfedési minta. A maximális hangnyomás színjének időfüggvényét folytonos görbével ábrázoltuk (bal ordináta skála). Pontokkal jelöltük a 2 s periódusú ismétlődő keskenysávú zajjal elfedett 3 kHz-es 2 ms időtartamú teszthang impulzus szintjének bizonyos késleltetési idők mellett mért értékeit (jobb ordináta skála)

4.5. Elfedési periódus minták

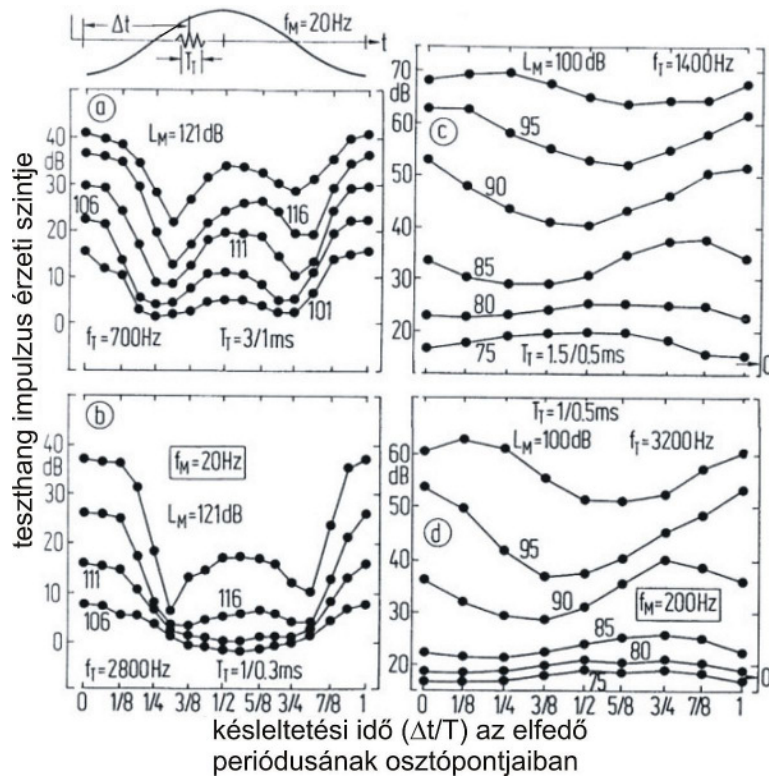
Az időbeli elfedési minta vizsgálatai során az elfedő jelnek csak a burkolóját tekintettük fontosnak, most az elfedési periódus minták (masking-period patterns) vizsgálatainál az elfedő hang időfüggvényét fogjuk felhasználni. Az elfedési periódus mintákban rövid, ma-



4.16. ábra. Alacsony frekvenciájú elfedési periódus minták (c) előállításához szükséges általános spektrális (a) és temporális (b) jellemzők. Az (a) részben egy 90 dB-es 100 Hz-es elfedő által előállított elfedési jelenség spektrumeloszlását jelöltük pontozott görbével az abszolút hallásküszöb (szaggatott vonal) felett. A vékony pontozott görbe a tesztjel spektrumeloszlását jelöli (1 ms időtartamú 2.5 kHz-es hangimpulzusok, a haranggörbe emelkedési-lecsengési ideje 0.5 ms; az ismétlési arány 100 Hz). Az ábra (b) részén a 10 ms-os periódus időbeli jellemzőit láthatjuk. A teszthang késleltetési idejét (Δt) az elfedő hang maximális ritkításának időpillanatától számítjuk. A teszthang impulzus periódikusan jelentkezik az elfedő egy periódusán belül, minden periódusban egyszer. A (c) részen látható elfedési periódus minta az elfedett teszthang impulzus szintjét ábrázolja az elfedő periódusának osztópontjaiban a késleltetési idő függvényeként. A paraméter az elfedő szintje

gas frekvenciájú teszthang impulzusok sorozatainak érzékelési szintjét vagy hangnyomás szintjét mérjük az elfedő teljes periódusán belüli helyük függvényeként. A teszthang impulzus frekvenciája néhány száz Hz-től néhány kHz-ig választható, de időtartamának rövidnek kell lennie az elfedő periódusához képest. A teszthang frekvencia felső határát akkor érjük el, ha az alacsony frekvenciájú elfedő hang már nem fedi el a teszthangot.

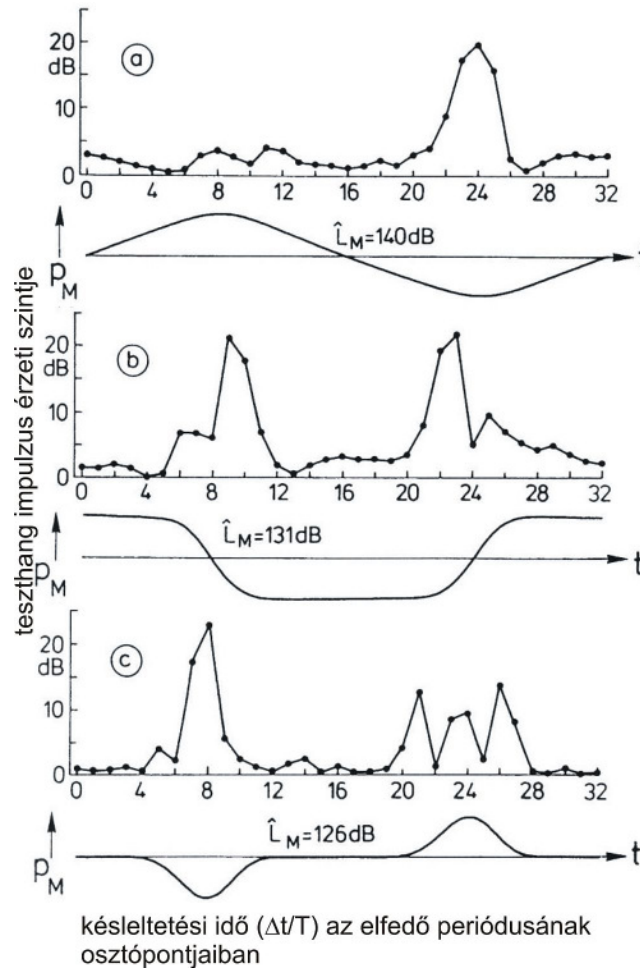
Az elfedő hang és a teszt jel spektrumeloszlásai és időfüggvényei adják a peremfeltételeket. A 4.16. ábra (a) része egy 90 dB-es 100 Hz-es szinuszos elfedő hang által előállított elfedés spektrumeloszlását mutatja. A nagy pontok azt a frekvenciatartományt jelölik, amelyben az elfedés előfordul. A szaggatott vonal az abszolút hallásküszöböt jelöli. A 2.5 kHz-es középfrekvenciájú, 1 ms-os időtartamú, 0.5 ms-os emelkedési-lecsengési idejű teszt hang impulzus sorozatok spektrum terjedelmét a kis pontokkal jelölt görbe ábrázolja. A görbék közül nyilvánvalóvá válik, hogy elfedési periódus mintát csak abban a frekvenciatartományban mérhetünk, amelyben az elfedő hang elfedési jelenséget eredményez. A második peremfeltétel az ábra (b) részén látható, ahol az elfedő hang és a tesztjel időbeli függvényeit ábrázoltuk. A tesztjelnek rövidnek kell lennie az elfedő periódusához képest, azaz rövidebbnek, mint az elfedő periódus nyolcada.



4.17. ábra. Elfedési periódus minták, azaz teszt hang impulzusok érzeti szintje késleltetési idejüknek függvényeként, 20 Hz-es (bal oldal) és 200 Hz-es elfedők esetén jelentkező elfedő periódusok osztópontjaiban. Minden panelen jelöltük a teszt hang frekvenciáját, időtartamát és a teszt hang impulzus emelkedési idejét. A paraméter az elfedő szintje

Ezen elfedő hangok és teszt hangok által előállított elfedő periódus minták láthatóak a 4.16. ábra (c) részén, ahol a paraméter az elfedő hang szintje. A 100 Hz-es elfedő hang esetén 75 dB alatt nincs elfedés. A 85 dB-es szint esetén az elfedés nagyobb az elfedő hang periódusának azon részén, ahol a periódusban a legkisebb értékeket elérjük (nevezzük ezt ritkulásnak, mivel a nyomás ilyenkor lesz kisebb, a levegő ilyenkor ritkább; ellentéte a sűrűsödés). A ritkulás maximuma referenciaként használható ($\Delta t = 0$) az időskálánál. A sűrűsödési maximum területén, ami a minta közepén van, az elfedési

küszöbök még alacsonyabbak lesznek, mint az abszolút hallásküszöb, amit a nyíl jelöl az ábra jobb ordinátáján. A 95 dB-es elfedési hang szinten az elfedési periódus minta elmozdul a teszthang impulzus magasabb szintjei felé (1/16-nál). 105 dB-en a minta még magasabbra mozdul el, és két csúcsot mutat az elfedő periódusán belül.



4.18. ábra. Elfedő hangok által előállított elfedési periódus minták. Minden panelen ábrázoltuk a hangnyomást és az időfüggvényt. Minden elfedő periódusa 300 ms. A (c) részen az elfedő időfüggvénye látható, a (b) részen annak első integrálját, az (a) részen a második integrálját ábrázoltuk. Minden panel felső részében egy 2 ms-os időtartamú, 0.5 ms-os emelkedési-lecsengési idejű, 1.4 kHz-es teszthang impulzus abszolút hallásküszöb feletti szintje látható a késleltetési idő függvényeként az elfedő periódusának 32 osztópontjában

Az elfedési periódus minta további jellemzőit mutatja a 4.17. ábra. Bal oldalon két rajzot láthatunk, ahol az elfedési hang frekvenciája 20 Hz, melynek periódusa 50 ms. A teszthang frekvenciája 700 Hz (felső ábrák) és 2800 Hz (alsó ábrák). A mintán belül a legmagasabb küszöböt mindig a ritkulás maximuma esetén érzük el. Néha van egy másik maximum is, de ez csak magasabb elfedési szinteken jelentkezik. Két minimumot láthatunk a két maximum között az elfedő hang időfüggvény zérusátmeneteinek közelében. Mivel a 200 Hz-es frekvencia csak 5 ms-os periódusnak felel meg, a teszthang impulzus

időtartamának elég rövidnek kell lennie.

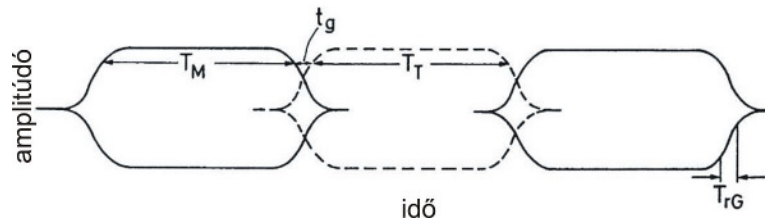
Hogy ezt a kapcsolatot még jobban jellemezhesük, nem-szinuszos időfüggvényű nagyon alacsony frekvenciájú elfedő hangokat használunk az elfedési periódus minták előállítására: váltakozó Gauss-görbéket, valamint azok első és második integrálját. Az elfedő hangok időfüggvénye a 4.18. ábrán látható a megfelelő elfedési periódus mintákkal együtt. Az elfedés periódusán belüli legalább 20 dB-es elfedés előállításához nagy elfedő szintekre van szükség. A kétszer integrált Gauss-görbe esetében (az ábra (a) része) a periódus első felében nincs elfedés, noha az elfedő szint 140 dB, viszont látható egy nagyon határozott maximum a második (ritka) felében. Valóban, a maximum megegyezik az elfedő hang időfüggvényének minimumával, azaz a ritkítás maximális értékével.

Az integrált Gauss-görbe esetén mért adatok (az ábra (b) része) 131 dB-es elfedő hang szinttel két határozott maximumot mutatnak annak ellenére, hogy ezeken a pozíciókon az elfedő hang időfüggvényének nincs szélsőértéke. A váltakozó Gauss-görbe (az ábra (c) része) olyan elfedési periódus mintát eredményez, amely maximumát a periódus első felében akkor éri el, amikor az elfedő hang időfüggvénye eléri a minimumát. Három maximum található még az elfedési periódus mintában, az elfedő hang periódusának második felében. A középső maximum ott található, ahol az elfedő hang időfüggvénye eléri legpozitívabb értékét. A két szomszédos maximumnál azonban nincs megfelelő szélsőérték az elfedő hang időfüggvényében. Az azonban nyilvánvaló, hogy az elfedő hang időfüggvényének második deriváltja játssza a fontosabb szerepet a periódikus alacsony frekvenciájú elfedő hangok által elfedett hangimpulzusok elfedési periódus mintáinak előállításában.

4.6. Pulzációs küszöb

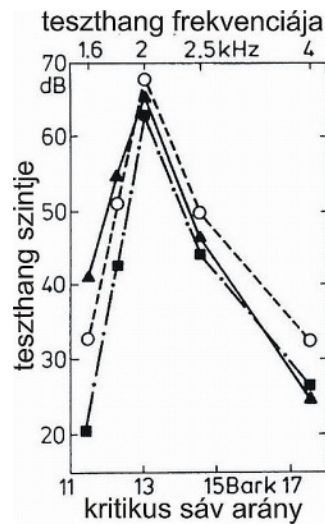
Ha két hang felváltva szólal meg, akkor a hangok egyikét folytonosként észlelhetjük annak ellenére, hogy valójában azt egyenletes időközönként be és ki kapcsoljuk. Egy ilyen érzékelés megfelelő időbeli mintáját mutatja a 4.19. ábra. A folyamatos görbék az „elfedőnek” nevezett hang időbeli burkolását, a szaggatott vonal pedig a teszthang időbeli burkolását ábrázolja. Az „elfedő” időtartamát (T_E -t), az „elfedő” és a teszthang közötti rést (t_g -t) és a teszthang időtartamát (T_T -t) a maximális amplitúdó 70%-án mérjük. A haranggörbe alakú kapuzó jel emelkedési és lecsengési idejét (T_{rG} -t) 10% és 90% között mérjük. Ha az „elfedő” és a teszthang sorozata elfogadható szinttel rendelkezik mindkét hang esetén, két pulzáló hang sorozatát észleljük. Azonban ha a teszthang szintje lecsökken egy meghatározott értékre, akkor a teszthangot folytonosnak kezdjük hallani, és nem pulzálónak. A teszthang szintnek ezt a sajátos értékét nevezzük pulzációs küszöbnek (pulsation threshold). Ha a teszthang szintje tovább csökken, a teszthangot továbbra is folyamatosnak halljuk, míg el nem éri az abszolút hallásküszöböt, mikor eltűnik. A pulzációs küszöb módszerben az „elfedő” valójában nem fedi el a teszthangot, viszont ez a hang kelti a folyamatosság érzetét. Ezért ebben a szakaszban az első hangra mindig „elfedőként” hivatkozunk.

Bizonyos értelemben, összehasonlítva a 3.2. szakaszban leírt elfedés által meghatározott elfedési mintákkal, a pulzációs küszöböket úgy is előállíthatjuk, hogy a pulzációs



4.19. ábra. A pulzációs küszöb mérésével kapott inger minták. Az ábrán az „elfedő” (folytonos) és a teszthang (szaggatott) időbeli burkolója látható; jelöltük az „elfedő” időtartamát (T_M), a teszthang időtartamát (T_T), az „elfedő” és a teszthang közötti rés időtartamát (t_g) és a haranggörbe alakú kapuzó jel emelkedési idejét (T_{rG})

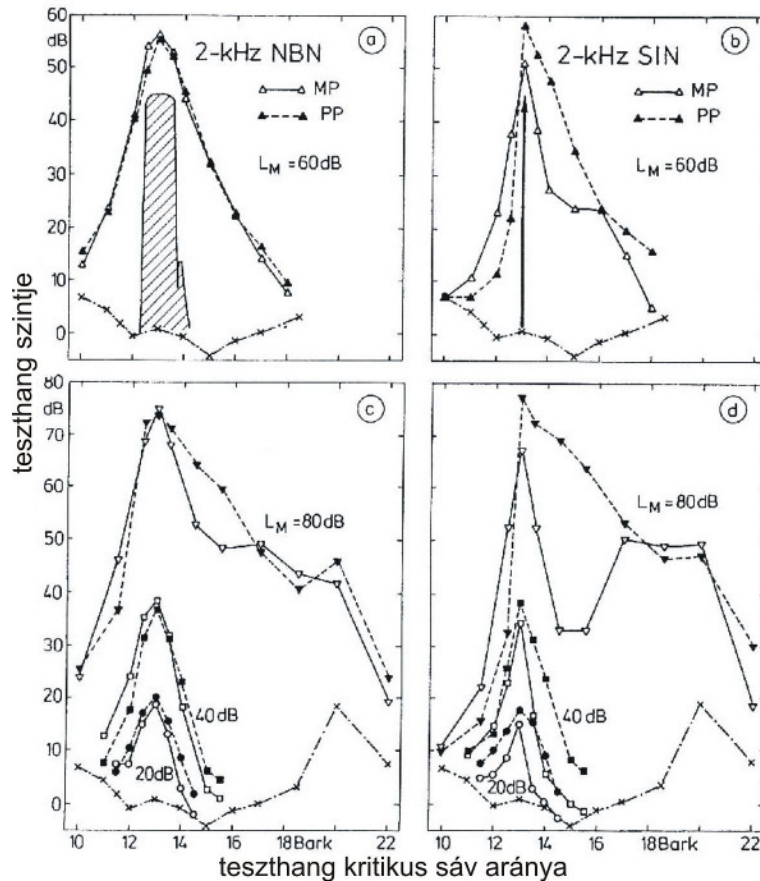
küszöb módszert alkalmazzuk. Egy 2 kHz-es kritikus sáv szélességű zajt „elfedőként” használva a kapott pulzációs mintát a 4.20. ábrán láthatjuk. A teszthang szintje (L_T) mind a teszthang frekvenciájának (felső skála), mind a teszthang kritikus sáv arányának (alsó skála) függvényeként adott. Míg az „elfedő” és a teszthang közötti rés időtartama (t_g), és a haranggörbe kapuzó jel emelkedési ideje konstans, az „elfedő” és a teszthang időtartama változó. Az „elfedő” szintje minden esetben konstans 70 dB.



4.20. ábra. Egy 2 kHz-es középfrekvenciájú kritikus sáv szélességű zaj „elfedő” pulzációs mintái. Az „elfedő” szintje 70 dB; az „elfedő” és a teszthang közötti időbeli rés 5 ms; a haranggörbe alakú kapuzó jel emelkedési-lecsengési ideje 10 ms; az „elfedő” és a teszthang időtartama 110 ms (háromszögek), 110 ms (körök) és 300 ms (négyzetek)

A 4.20. ábrán szereplő eredmények azt mutatják, hogy a pulzációs minta alakja kevésbé függ az „elfedő” és a teszthang időtartamának választásától. A pulzációs minta alacsonyabb frekvenciáinak meredeksége csökken némileg az „elfedő” és a teszthang időtartamának csökkenésével. Megpróbálhatunk egy „optimális” inger mintát kifejleszteni a pulzációs küszöbök mérésére. Az argumentumok a 4.21. ábrán bemutatott eredményektől függenek, ahol a bal oldali ábrák egy 2 kHz-es (13 Bark) középfrekvenciájú keskenysávú zaj esetén, a jobb oldali ábrák egy 2 kHz-es tiszta hang esetén kapott eredményeket mutatják.

Az elfedési mintákat üres háromszögek és folytonos vonalak, a pulzációs mintákat tele háromszögek és szaggatott vonalak jelzik.

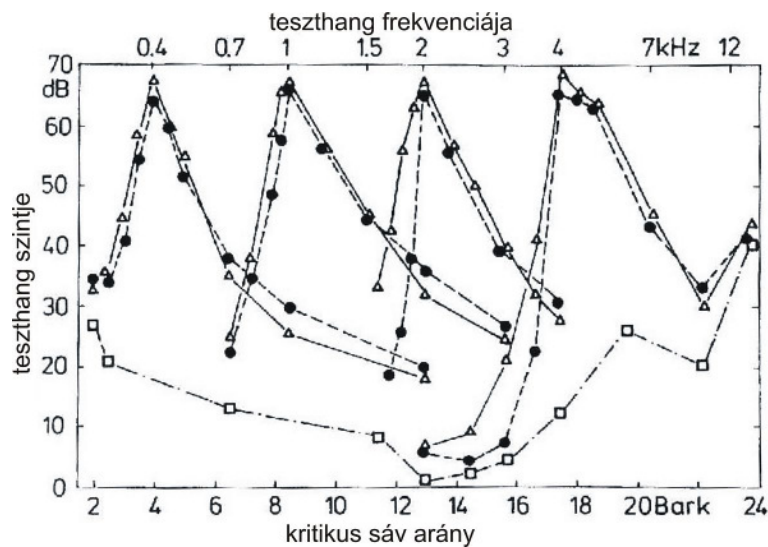


4.21. ábra. „Elfedők” által előállított elfedési minták és pulzációs minták összehasonlítása: bal panelek: 2 kHz-es (13 Bark) középfrekvenciájú kritikus sáv szélességű zaj; jobb panelek: 2 kHz-es tiszta hang. Az üres háromszögek és folytonos vonalak az elfedési mintákat, a tele háromszögek és szaggatott vonalak a pulzációs mintákat, a keresztek és szaggatott-pontozott vonalak az abszolút hallásküszöböt jelölik. Az (a) rész a 60 dB hangnyomású kritikus sáv szélességű zajt (a vonalkázott terület a kritikus sáv szélességű zaj spektrumeloszlását jelöli), a (b) rész a 60 dB hangnyomású tiszta hangot (a nyíl az elfedő hang spektrális pozícióját jelöli), a (c) és (d) rész a 20, 40 és 80 dB hangnyomású kritikus sáv szélességű zajokat és tiszta hangokat ábrázolja

Míg az elfedési minták mérései keskenysávú zaj esetén egyszerűek, a tiszta hang elfedési mintáit különféle hatások zavarják meg (lásd a 3.4. szakaszt). Ezért a stratégia a következő: egy 60 dB szintű 2 kHz-es kritikus sáv szélességű zaj esetén az elfedési mintát és a pulzációs mintát is meghatározzuk. Aztán változtatjuk az inger temporális jellemzőit a pulzációs küszöb 4.19. ábrán látható vizsgálata esetén, amíg az eredményül kapott pulzációs mintát olyan közel nem illesztjük a megfelelő elfedési mintához, amilyen közel csak lehetséges. A 4.21. ábra (a) részén látható, hogy az elfedési minta és a pulzációs minta nagyon közel lehet egymáshoz, az átlagos különbség a két minta között kevesebb,

mint 1 dB. Ez viszont csak akkor teljesül, ha az inger jellemzői a következők: az elfedő hang időtartama, $T_E = 100$ ms, a teszt hang időtartama, $T_T = 100$ ms, a temporális rés az elfedő hang és a teszt hang között, $t_g = 5$ ms és a $T_{rG} = 10$ ms. Ebben a szakaszban a továbbiakban ezt az „optimális” inger mintát fogjuk használni a pulzációs minta mérésére. Hasonlítsuk össze a folytonos és szaggatott vonalakat az ábra (c) részén. Ez azt mutatja, hogy nemcsak 60 dB-es szint esetén, hanem 20, 40 és 80 dB mellett is eléggé hasonlóak a keskenysávú zaj pulzációs mintái és elfedési mintái. Csak a 80 dB-es görbe magasabb frekvenciáin mutatkozik különbség az elfedési görbe nemlinearitása miatt, ami a pulzációs mintában nincs jelen.

A jobb oldali ábrákon hasonló összehasonlítás látható, ha az elfedő hang szinuszos tiszta hang. A keskenysávú zaj esetén látott eredményekkel ellentétben itt sokkal nagyobb különbségek jelentkeznek az elfedési minta és a pulzációs minta között. Ez a különbség akkor is jelentős, ha ugyanazt a konfigurációt használjuk, mint keskeny sávú zaj esetén. Azonban a pulzációs minták ugyanazt a „dőlést” mutatják, mint amit a tiszta hangok elfedési mintáiban láthattunk: alacsony elfedő hang szint mellett a minta meredeksége az alacsonyabb frekvenciákon kisebb, mint magasabb frekvenciákon, a magas elfedő hang szinteken pedig az alacsonyabb frekvenciákon nagyobb a meredekség.



4.22. ábra. Kritikus sáv szélességű zaj és tiszta hang pulzációs mintáinak összehasonlítása. Az „elfedő” szintje 70 dB. Az „elfedő” frekvenciája 0.4, 1, 2 és 4 kHz; a háromszögek és folytonos vonalak a kritikus sáv szélességű „elfedőt”, a körök és szaggatott vonalak a tiszta hang „elfedőt”, a négyzetek és szaggatott-pontozott vonalak a nyolc megfigyelésből származó abszolút hallásküszöbök középvonalát jelölik

Ha összehasonlítjuk a 4.21. ábra (a) és (b) részét, akkor megállapíthatjuk, hogy a tiszta hang esetén látható minták az alacsonyabb frekvenciákon meredekebbek, mint a keskenysávú zaj esetén előállt minták. Ez különösen igaz a pulzációs mintára. Vizsgáljuk meg, előfordul-e ez a jelenség más középfrekvenciákon is. A 4.22. ábrán a 0.4, 1, 2 és 4 kHz-es elfedő hangok mellett előálló pulzációs minták láthatóak. A háromszögek és folytonos vonalak a kritikus sáv szélességű zaj esetén, a körök és szaggatott vonalak a tiszta hang

esetén kapott eredményeket reprezentálják. A lent látható pontozott szaggatott vonallal összekötött négyzetek egy nyolc vizsgálatból származó hallásküszöb középértékét jelölik. A kritikus sáv szélességű zajok esetében és a tiszta hangok esetében is 70 dB a hangnyomás szint.

Ha összehasonlítjuk a pulzációs mintákat az alacsonyabb frekvenciákon, akkor azt látjuk, hogy 0.4 kHz és 1 kHz esetén nincs lényegi különbség. Azonban 2 kHz esetén már nagy, 4 kHz esetén ettől valamivel kisebb a különbség tiszta hang és keskenysávú zaj esetén. A 4.22. ábrán látható pulzációs minták nagyjából azonosak a kritikus sáv szélességű zaj 3.3. ábrán látható elfedési mintáival. Ez a hasonlóság viszont csak akkor igaz, ha $T_E = T_T = 100ms$ és $t_g = 5ms$.

4.7. Kevert spektrális és temporális elfedés

Sok esetben, ha összetett elfedésről beszélünk, a spektrális és temporális hatások egyaránt jelen vannak. Azonban ezekben az esetekben túl sok a változó, így ezt a jelenséget nem tárgyaljuk részletesen, csak három példán keresztül mutatjuk be.

Az elfedés alapja mindig ugyanaz: a gerjesztési szint, kritikus sáv arány, az elfedő hang és a teszthang által előállított időminták közötti (spektrális és temporális) arány. Ez a felelős azért, hogy a teszthang hallható-e vagy sem.

Egy frekvenciamodulált tiszta elfedő hang konstans hangnyomás szintet produkál. Azonban a gerjesztési szintje fel-le mozog a kritikus sáv skála mentén egy, a frekvencia ingadozásának megfelelő terjedelemmel és egy, a modulációs frekvenciájának megfelelő időtől függő sebességgel. Ha teszthangként egy hosszan tartó hangot használunk, a moduláció periódusán belül előálló elfedő hang minimális gerjesztési szintje a teszthang kritikus sáv arányától függő küszöböt határoz meg. Ha egy rövid teszthang impulzust használunk, ami a modulációs frekvenciának megfelelő mértékben többször ismétlődik, akkor az elfedő hang frekvenciamodulációja és a frekvenciamoduláció periódusán belül jelen lévő impulzus előfordulása közötti időbeli kapcsolat járulékos szerepet játszik, ahogy azt a 4.9. ábrán is láthattuk.

Azonban FM-szerű hangokat előállíthatunk úgy is, ha egy összetett elfedő hang nagy számú tonális komponenseinek fázisait speciális szabályokat követve határozzuk meg. Ha a fázisokat véletlenszerű értékekről erre a speciális feltételre változtatjuk anélkül, hogy a részleges hangok szintjét megváltoztatnánk, hallásunk fázis érzékenységét meghatározhatjuk.

A Gauss zajt alacsony szintű, véletlenszerű fázissal rendelkező szinuszos komponensek segítségével szimulálhatjuk. Ezt azonban csak akkor fogadhatjuk el jó közelítésnek, ha elfedő hangként csak óvatosan használjuk. Ugyanez igaz azokra a zajokra is, amelyeket „0”-ások és „1”-esek véletlenszerű sorozata állít elő. Ezek a zajok gyakran a Gauss zaj által előállított elfedési jelenségektől eltérő elfedési jelenségeket eredményeznek az eltérő amplitúdó eloszlás és a keskenyfrekvenciájú sávokban lévő eltérő temporális és spektrális minták miatt. Ezek a minták fontosak a gerjesztési szint esetén, a kritikus sáv arány esetén és az időminták esetén, tehát az elfedés alapjainál.

5. fejezet

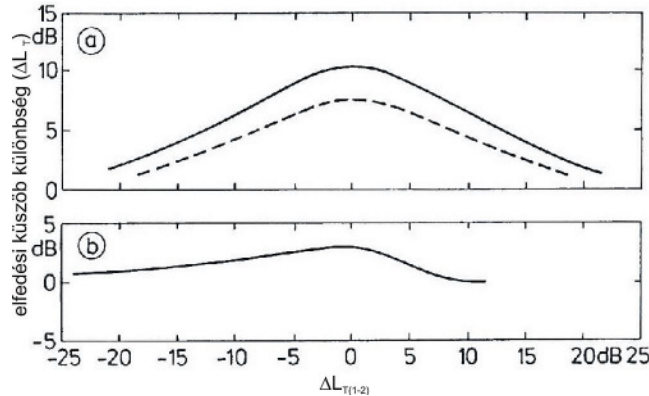
Az elfedés „kiegészítése”

A 3. fejezetben leírtuk az egyes jól meghatározott elfedő hangok által produkált elfedési jelenségeket. Sok esetben azonban nemcsak egy, hanem egy második elfedő hang is jelen van, ami lehet például egy 1 kHz-es középfrekvenciájú keskenysávú zaj. Az alacsony frekvenciák esetén az alacsony frekvenciájú elfedő hang fogja meghatározni az elfedést. Az 1 kHz körüli frekvenciákon az elfedést a keskenysávú zaj határozza meg. Van azonban egy terület, amelynél a keskenysávú zaj által és az alacsony frekvenciájú hang által produkált elfedési hatás külön jelentkezik egy időben. Ez a küszöb egy egyenlő mértékű eltolását eredményezi. Ez az említett esetben valószínűleg 800 Hz közelében lesz. Azt a jelenséget, amelyre az elfedés „kiegészítése”-ként hivatkozunk, a következő kérdéssel lehet megfogalmazni: a két zaj által elfedett küszöb ugyanaz marad-e, mint amikor a küszöböt csak egy hanggal fedjük el, vagy esetleg 3 dB-el (intenzitás kiegészítés), 6 dB-el (hangnyomás kiegészítés) vagy más dB értékkel emelkedik? A szimultán elfedést használó vizsgálatok megfelelő eredményei viszonylag könnyen összegezhetőek. A nemszimultán elfedés (pl. utóelfedés) használatával kapott hasonló eredmények azonban olyan jelenségeket eredményeznek, amelyeket nem lehet könnyen általánosítani. Úgy tűnik, hogy a sáv szélesség más paraméterekkel együtt fontos szerepet játszik ebben az esetben.

5.1. A szimultán elfedés „kiegészítése”

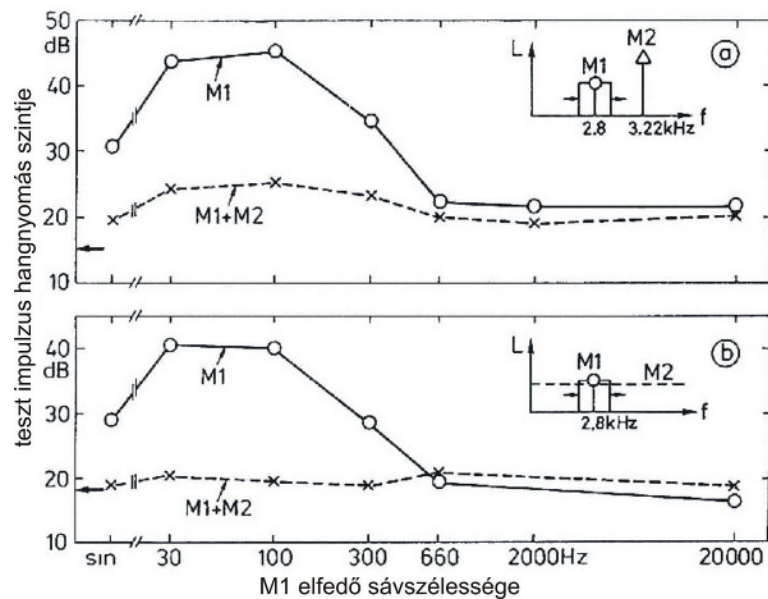
Az elfedő hangok megkülönböztetésére egy indexet használunk. Például ha ki akarjuk jelezni az elfedési küszöbök szintjét, akkor azokat jelölhetjük L_1 -el és L_2 -vel, és a második elfedő hang hozzáadásával kapott elfedési küszöb hozadékát a növekmény értékével, ΔL_T -vel. A legfontosabb változók egyike a két elfedő hang által produkált két küszöb (L_1 és L_2) közötti különbség. Ezt a különbséget jelöljük $\Delta L_{T(1-2)}$ -vel. Azt várjuk, hogy a küszöb növekedése, ΔL_T akkor lesz a legnagyobb, ha az elfedő hangok által előállított elfedési küszöbök közti különbség nulla, azaz ha mindkét elfedő hang ugyanazt az elfedési küszöböt eredményezi. Ha az egyik elfedő hang sokkal nagyobb küszöböt produkál, mint a másik, akkor az együttesen előállított elfedés során a nagyobb küszöböt eredményező elfedő hang lesz a domináns.

A két elfedő hang jelenléte mellett létrejövő elfedés ΔL_T különbsége látható az 5.1. ábrán, ahol összehasonlíthatjuk azt az egyetlen elfedő hang eredményével. Két görbét



5.1. ábra. A teszthangokat vagy két elfedő (egy 105 Hz-es hang és egy 1 kHz középfrekvenciájú kritikus sáv szélességű zaj) együttesen vagy a nagyobb elfedést produkáló elfedő egyedül fedi el. A teszthangok elfedési küszöbeinek különbségét (ΔL_T) a két elfedő egyenkénti alkalmazása által előállított elfedési küszöbök különbségeinek ($\Delta L_{T(1-2)}$) függvényeként ábrázoltuk. Az (a) rész az elfedési minták meredekségein (emelkedésén és lejtésén) mérhető adatokat (folyamatos: átlaggörbe 15 dB-el az abszolút hallásküszöb felett lévő elfedési küszöbök esetén, szaggatott: abszolút hallásküszöb felett 5 - 15 dB-el) mutatja. A (b) rész a zaj elfedő kritikus sávján belül előálló adatokat mutatja

láthatunk az ábrán, a szaggatott vonalú azoknak az elfedési küszöböknek felel meg, amelyek csak 5 - 15 dB-el vannak az abszolút hallásküszöb fölött, a folytonos görbe a több mint 15 dB-el az abszolút hallásküszöb felett lévő elfedési küszöböknek felel meg. A szintkülönbségnél, $\Delta L_{T(1-2)} = 0$ -nál várt különbség tisztán látható mindkét esetben. Azonban az elért maximális ΔL_T érték sokkal nagyobb mint 3 dB, és nagyon közel a 2-es tényezőhöz a 6 dB-es értéket is meghaladja. Az abszolút hallásküszöbnél jóval nagyobb esetek vizsgálatai során gyakran 12 dB-t kapunk $\Delta L_{T(1-2)} = 0$ esetén. A görbe csökkenése $\Delta L_{T(1-2)}$ negatív és pozitív értékei felé a vártnál kisebb meredekségű. Az elfedési küszöbök növekedése akkor is eléri a 6 - 8 dB-t, ha a mindkét elfedő hang által produkált elfedési küszöb közti különbség 10 dB nagyságrendű. Az ábra (a) részének adatai azokra az elfedési küszöbökre vonatkoznak, amelyeket a gerjesztés meredekségein határoztak meg. Ez azt jelenti, hogy az elfedési küszöböt nem az elfedő hang kritikus sáv szintjén, azaz a fő gerjesztésen határozzuk meg, hanem a magasabb vagy alacsonyabb gerjesztési meredekségeken. Ha a méréseket egy fő gerjesztésként meghatározott frekvencia tartományban hajtjuk végre (a mi esetünkben ez a keskenysávú elfedő hang középfrekvenciájának, 1 kHz-nek a környéke), az elfedési küszöbökben elért növekedés csak 3 dB lesz, ha a két elfedő hang ugyanazt az elfedést produkálja (lásd az ábra (b) részét). Ez azt jelenti, hogy ilyen esetekben ezen elfedő hangok elfedési kiegészítése összefüggésben van az intenzitási kiegészítéssel.



5.2. ábra. Utóelfedési küszöbök olyan teszthang impulzusok esetén, melyeknek frekvenciája az M1 elfedő középfrekvenciája. A folyamos vonalak azokat a küszöböket jelölik, amelyeket az M1 egyedül állít elő, a szaggatott vonalak pedig az M1 és M2 egyidejű alkalmazása mellett előálló küszöböket. A nyilak az M2 egyedüli alkalmazása mellett előálló utóelfedési küszöböket jelölik. Az M1 kritikus sáv szélességű elfedő középfrekvenciája 2.8 kHz, szintje 40 dB. Az M2 3.22 kHz-es sáv szélességű és 60 dB hangnyomású tiszta hang az (a) részen, és szélessávú zaj 20 kHz-es sáv szélességgel és 60 dB hangnyomás szinttel a (b) részen

5.2. Az utóelfedés „kiegészítése”

Ahogy már korábban is tárgyaltuk, az utóelfedés függ az elfedő hang időtartamától. Ha keskenysávú elfedő hangokat használunk, akkor az utóelfedés a sáv szélességtől is függ. Ezért sok más változó mellett a két elfedő hang szintje is befolyásolja az „1”-es és „2”-es elfedő hang által létrehozott jelenség kiegészítését. Nézzünk meg egy példát, ami viszonylag jelentős hatásokat produkál. Mindkét elfedő hang legyen relatíve hosszú (körülbelül 500 ms). A teszthang időtartama 20 ms, alakja haranggörbe, 2 ms-os emelkedési idővel. A késleltetési idő az elfedő hang vége és a teszthang vége között konstans 46 ms. Egy 2.8 kHz-es középfrekvenciájú keskenysávú zaj elfedő hang jele legyen M1. A teszthang impulzus frekvenciája ugyanaz, mint M1 középfrekvenciája. Az M2 elfedő hang egy 3.22 kHz-es szinuszos hang (5.2. ábra (a) része) vagy egy szélessávú zaj (az ábra (b) része). A spektrumok konfigurációját az ábrákban lévő rajzokon láthatjuk. Ezeket az elfedési küszöböket mindkét esetben az M1 elfedő hang sáv szélességének függvényeként ábrázolják. A görbék nyolc vizsgált személy 16 adatpontjának középértékét ábrázolják. A baloldalon látható nyilak az M2 elfedő hang egyedüli alkalmazása mellett észlelhető utóelfedési küszöböt jelölik. Mindkét ábra legbaloldali adatpontjai egy M1 szinuszos elfedő hanghoz tartoznak. Ettől az elfedő hangtól továbbhaladva egy 30 Hz-es sáv szélességű keskenysávú elfedő hang felé körülbelül 12 dB-es elfedési küszöbnövekedést kapunk. A 100 Hz-es sáv szélesség fölött csökkenni kezdenek egy majdnem-konstans értékhez, melyet 660 Hz-nél érünk el. Ha a két

elfedési hang szimultán jelentkezik, akkor az utóelfedési küszöbök majdnem függetlenek maradnak az M1 elfedő hang sávszélességétől. Az utóelfedési küszöbök nagymértékű csökkenéséhez vezet, ha a második elfedő hangot, M2-t hozzáadjuk. Ez a jelenség teljesen elletétes a szimultán elfedés során látottakkal. Utóelfedés esetén a második elfedő hang hozzáadásával kapott elfedés csökkenése erősen függ a paraméterek választásától.

Az 5.2. ábrán látható példák a legjellegzetesebb esetet mutatják. Az adatok szerint ez a meglepő hatás függ M1 elfedő hang időstruktúrájától, amely jellegzetes időmintát eredményez 600 Hz alatti keskeny sávszélességeken. Ezért azoknak a modelleknek, amelyek ezeket az eredményeket leírják, a két elfedő hang spektrális és temporális jellemzőire kell épülniük. Így az M1 keskenysávú elfedő hang esetén jelentkező egyértelműen hallható ingadozásokat is figyelembe veszi. Ezek az ingadozások redukálódnak, ha egy második elfedő hang, M2 is csatlakozik, amely lehet egy tiszta hang vagy egy fehér zaj. Végül a spektrális hatásokon felül egy jellegzetes időbeli hatás a hallható ingadozások csökkenéséhez vezet, mely végül az utóelfedési küszöb 13 - 24 dB mértékű csökkenését eredményezi. Az utóbbi jelenség erősen függ az elfedő hang sávszélességétől. Ha M1 keskenysávú zaj, akkor jelentős hatások jelentkeznek, míg ha szélessávú zaj, akkor szinte észrevehetetlen a hatás.

6. fejezet

A szimultán elfedés pszichoakusztikai modellje

A 200 ms-nál hosszabb éppen hallható hangváltozások modellje azt feltételezi, hogy egy változás olyan feltételek mellett hallható, amelyben a gerjesztési szint 1 dB-nél nagyobb mértékben növekszik a kritikus sáv arány skála bármely területén. Mivel az elfedés és a gerjesztés közeli kapcsolatban vannak, a növekménynek tartalmaznia kell a tartós állapot esetén jelentkező elfedési jelenségeket is, mint például a tiszta hang hosszantartó szélessávú zajjal történő elfedését. A zajok ingadozását, amely kis sáv szélességekkel történik, és a hallórendszer kezdeti szakaszában áll elő annak frekvencia szelektivitása miatt, szintén figyelembe kell venni. Alacsony frekvenciákon például a megfelelő kritikus sáv szélesség csak 100 Hz, ugyanakkor magas frekvenciákon, 10 kHz környékén a sáv szélesség több mint 2 kHz. Következésképpen a magas frekvenciájú zaj tartós állapotot eredményez. Viszont az alacsony frekvenciájú zajok erős temporális modulációt állítanak elő, ami megakadályozza a hallás folyamatát, ha az elfedési eljárásban megszakított teszthangot használunk.

A modell használatát egy példával lehet illusztrálni, ahol az elfedő hang fehér zaj, a teszt hang pedig egy szinuszos hang. A fehér zaj egy szélessávú zaj, ami csak fő gerjesztéseket állít elő. Magas frekvenciákon a zaj ingadozásának nincs hatása, mert a kritikus sáv szélesség nagyon nagy. Ezért a logaritmikus küszöb együttható 1 dB, az éppen hallható relatív növekedés az intenzitásban 0.25, és a megfelelő elfedési index -6 dB. Tegyük fel, hogy a teszthang frekvenciája 9 kHz, a megfelelő kritikus sáv szélesség 2000 Hz. Mivel a fő gerjesztési szint megegyezik a kritikus sáv szinttel, az elfedési küszöböt az $L_T = L_G + a_v = l_{WN} + 10 \log(f_G/Hz) + a_v$ egyenlőséggel számíthatjuk ki. A fenti példa és egy $l_{WN} = 30$ dB-es elfedő hang sűrűségi szint esetén a teszthang szintjét az elfedési küszöbnél a következőképpen számíthatjuk ki: $L_T = (30 - 10 \log 2000 - 6) \text{ dB} = (30 + 33 - 6) \text{ dB} = 57 \text{ dB}$. Ez az érték egyezik a 3.1. ábrán látható adatokkal.

Hasonló számítások végezhetők az alacsony frekvenciák esetén. Itt az elfedési index csak -2 dB, az éppen hallható relatív intenzitásnövekedés 0.65, és a magas frekvenciák esetén jelentkező 1 dB-es növekedés logaritmikus értéke csak egy kicsit másabb, mint egy alacsony frekvenciák esetén jelentkező 2 dB-es növekedés. Például 300 Hz-es teszthang frekvencia mellett a kritikus sáv szélesség 100 Hz, a teszthang szintje az elfedési küszöbnél $L_T = (30 - 20 - 2) \text{ dB} = 8 \text{ dB}$. Ez az érték szintén egyezik a mért adatokkal. Az

egyezés nem meglepő, mivel a gerjesztési szint és kritikus sáv arány minta az elfedési küszöb mintából állt elő, és az elfedési modell ezt az összefüggést használja az ellenkező irányban.

Az abszolút hallásküszöb közelében az elfedési küszöb függetlenné válik nemcsak az elfedő hang szintjétől, de az abszolút hallásküszöb szintjétől is. Ebben az esetben az elfedő zajt és az abszolút hallásküszöbért felelős belső zajt össze kell adni. Ilyen feltételek mellett az alacsony szintek szélessávú zajával elfedett hangok küszöbe nagyon egyenletesen alakul át elfedési küszöbből abszolút hallásküszöbbe. A modellbe tehát felvehetjük az abszolút hallásküszöböt, ha elfedő hangként egy belső zajt tekintük.

A szimultán elfedés függ a teszthang időtartamától oly módon, hogy az intenzitás és az időtartam szorzata konstans a 200 ms-nál rövidebb időtartamok mellett. Hosszabb időtartamok esetén a küszöb függetlenné válik az időtartamtól. Ez a jelenség úgy is leírható, mint egy 200 ms-os jellemző értékkel rendelkező időkonstanssal terhelt intenzitás. Ha szélessávú zaj helyett keskenysávú zajt vagy hangot használunk elfedő hangként, gerjesztési meredekség jelentkezik. Mivel a gerjesztési meredekség keskenysávú hangok által előállított elfedési mintákon alapul, a fent leírt tények még jobban alkalmazhatóak: az elfedés modellje a gerjesztésen és a kritikus sáv arányon alapszik, ezért nyilvánvaló, hogy ez a modell tökéletesen leírja a tartós állapotú spektrális elfedést.

7. fejezet

Hangtömörítés

Az audio tulajdonképpen hangnyomásváltozásoknak tekinthető, amely változások folyamatosak időben és amplitúdóban. Az ilyen jelek számítógépen való tárolásához diszkrétizálásra van szükség az időtartományban és az amplitúdó tartományban egyaránt, ezt nevezzük mintavételezésnek. A mintavételezés során a folyamatos jelet a mintavételi időpontokban a folyamatos jel értékeivel megegyező értékek sorozatával helyettesítjük. A mintavételi időpillanatok egymástól egyenlő távolságra helyezkednek el, ezt az időtartamot mintavételi időnek, a reciprokát pedig mintavételi frekvenciának nevezzük.

Az emberi hallás a 20 Hz és 20 kHz közötti frekvenciákat észleli. 44.1 kHz mintavételezési frekvenciát használnak az olyan forgalomban lévő digitális HIFI termékek, mint a CD. A professzionális alkalmazások tipikusan 48 kHz-es mintavételt használnak.

A mintavett jelet végessé és diszkrétté kell tenni, amit a kvantálással valósítunk meg. A kvantálás tehát a jel értékkészletének diszkrétté tétele. Ekkor a véges sok érték közül egyet kiválasztunk valamilyen módszerrel, és azt használjuk a továbbiakban. Általában 16 bites kvantálást alkalmaznak, ami 65536 kvantálási érték megkülönböztetését jelenti. A kvantálási értékek elemszámának kettes alapú logaritmusát bitsebességnek nevezzük. Ha a kvantálás nagyon durva (keves bitet használunk), a rekonstruált audio jel másképp fog hangzani, mint az eredeti. Azt a zajt, amit a kvantálás állít elő, kvantálási zajnak nevezzük, ez határozza meg az elérhető legnagyobb jel-zaj arányt. A jel-zaj arány decibelben van kifejezve, és definíciója:

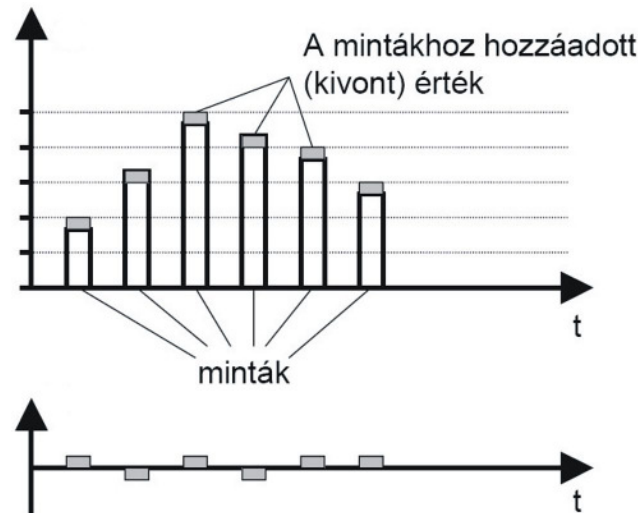
$$S/N = 10 \log(V_{signal}^2 / V_{noise}^2) = 20 \log(V_{signal} / V_{noise}). \quad (7.1)$$

A CD-k 16 bittel vannak rögzítve, és 96 dB-es maximális jel-zaj arányt érnek el. A professzionális alkalmazások 18 vagy 20 bitet használnak, ami még magasabb jel-zaj arányt eredményez. A minta értékeket tipikusan lineáris vagy logaritmikus formában tárolják (ezek neve μ -törvényű és A-törvényű). Az utóbbi növeli a lehetséges dinamikus tartományt, de több zajt eredményez.

Azt a folyamatot, amit eddig leírtunk, PCM (egyenletes kvantáló; pulse-code modulation) kódolásnak nevezzük. A kapott eredményt különböző kódolókkal tovább tömöríthetjük.

A dekódoló a kódoló hatását alakítja vissza. A kvantálási zaj dekódoláskor bennmarad a jelben. Ez a hiba magasfrekvenciájú zajként jelenik meg, ezért a visszaalakítás során

aluláteresztő szűrőt használunk, amely a zaj nagy részét kiszűri a jelből. A kvantálás és a kvantálási zaj a következőképpen néz ki:



A felső ábrán a kvantálás folyamata, az alsón pedig a folyamat során előállt kvantálási zajt láthatjuk.

A tömörítési eljárás akkor lesz jó hatásfokú, ha az elfedés jelenségét maximálisan kihasználja. Az elfedés mellett fontos szerepük van a kritikus sávoknak is. A kritikus sávokat a 2.2. szakaszban tárgyaltuk. Láthattuk, hogy a hallható frekvenciatartomány lefedhető 24 kritikus sávval. A kódolók, amelyek felhasználják a kritikus sávok létezését is, az audio jelet a frekvencia tartományba transzformálják, majd az így kapott spektrumot felosztják a kritikus sávoknak megfelelően, és végül az egyes alsávokat a kvantálási zaj hallhatóságának függvényében kvantálják.

Optimális tömörítés esetén minden egyes sávhoz éppen akkora kvantálási szintet kell biztosítani, amely feltétlenül szükséges ahhoz, hogy a kvantálási zaj hallhatatlan legyen.

7.1. Szokásos tömörítési eljárások

A tipikus veszteségmentes tömörítési eljárások felismerik az adatok ismételt sorozatait, és a gyakran előforduló szimbólumokat rövidebb szimbólumokként statisztikailag rögzítik. Ez jól működik szövegeknél, számítógépes programoknál és más adatfolyamoknál, amelyek nagy mértékű redundanciát tartalmaznak. Azonban az olyan adatoknál, mint a hang és a képek, ez a tömörítési módszer egyáltalán nem lesz hatásos. Ezért sok veszteséges tömörítési eljárást terveztek digitális hangok számára, amelyek változó mértékű hallható veszteséggel dolgoznak.

A veszteséges tömörítést megvalósító eljárások, nem teszik lehetővé az adatok pontos rekonstrukcióját, ám a rekonstrukció elég jó ahhoz, hogy ne következzen be érzékelhető minőségromlás. A veszteséges hangtömörítés megvalósítására többféle szabvány létezik. A továbbiakban ezeket tárgyaljuk.

A veszteséges tömörítési eljárások két csoportra oszthatók, az egyik csoport az időtartományon, a másik a frekvenciatartományon dolgozik.

Az egyik szokásos eljárás, hogy kiszámoljuk és kvantáljuk az egymást követő minták, vagy a következő minta és a következő minta becslése közötti különbséget. Remélhetőleg ez a különbség kisebb és kvantálhatóbb lesz, mint az eredeti minta. Ennek a kódolási eljárásnak DPCM (Differential Pulse Code Modulation) a neve. Ez továbbfejleszthető úgy, hogy alkalmazzuk az előrejelzést az aktuális jel jellemzőire, ezt a módszert ADPCM-nek (Adaptive Differential Pulse Code Modulation) nevezzük, és ez a beszédet 16 és 40 kbps-al tudja tömöríteni a minőségi kívánalmaktól függően.

Van egy másik módszer is a hang modellezésére, amikor a hangot matematikai modellre képezzük le, aztán csak transzformáljuk a modellenél szereplő tényezőket az aktuális minta értékek helyett. Ezeket a tömörítőket vokódereknek hívjuk, erre példa az LPC (Linear Predictive Coding). Ez csak 2.4 kbps-ot eredményez, de túlságosan veszteséges, és nagyon szintetikusnak halljuk. Az LPC egyik továbbfejlesztése a CELP (Code Excited Linear Predictor), ami úgy működik, mint az LPC, de megpróbálja az adatokat - amennyire lehetséges - az adott bitsebesség alatt továbbítani. A 4.8 kbps-os bitsebességű CELP-et használva tűrhető telefonvonal-minőségű hangot kapunk.

A részsávú kódolás. A részsávú kódolás (sub-band coding - SBC) az elfedéstől függ. Az egészséges fül sok frekvenciát tud érzékelni. Azonban ha egy meghatározott frekvencián sok az energia, akkor a fül nem képes érzékelni a közeli frekvenciák alacsonyabb energiáit. Tehát az erősebb frekvencia elfedi a gyengébb frekvenciákat. Ez az elfedés egész pontosan a szimultán elfedés.

A részsávú kódolás egymást érő frekvenciasávokra osztja a spektrumot. A kód csak azoknak a részsávoknak az értékeit tartja meg, amelyek nincsenek elfedve.

Először egy gyors Fourier-transzformációval (FFT) felbontjuk a jelet részsávokra. A pszichoakusztikus modell meghatározza ezen részsávok elfedési küszöbeit a pszichoakusztikai információk felhasználásával. A részsávokban vizsgáljuk, hogy milyen energiájú jelek vannak a küszöb alatt, ezeket egyszerre tekintjük, egyetlen komponensként kezeljük őket. A kvantálást úgy megvalósítani, hogy a kvantálási zaj az elfedési küszöb alatt maradjon. Végül a kvantált mintákat keretekbe rakjuk. A dekódolásnál ezeket a csomagokat csak ki kell bontani, és az inverz FFT-vel visszaalakítani audió jelekké.

Sok különböző frekvenciatartomány alapú tömörítési szabvány van, de a legsikeresebbek ugyanazon az alapelven osztoznak, amit pszichoakusztikának nevezünk.

Pszichoakusztika. Az emberi hallórendszer 96 dB feletti dinamikus tartománnyal rendelkezik. Azonban nyilvánvaló, hogy amíg nagyon halk hangokat, mint egy tű leesése, és nagyon hangos zajokat, mint a repülőgép felszállása könnyen meghalljuk, képtelenek vagyunk meghallani a leeső tűt, ha ugyanakkor a repülőgépet is halljuk. Az emberi hallás alkalmazkodik a hang dinamikus változásaihoz, és ezek az alkalmazkodások és elfedési hatások alkotják a pszichoakusztikai elméletek és kutatások alapját.

Az elfedés térbeli és időbeli környezetében a zajt vagy a halk hangokat nem tudjuk észlelni, ha azok a küszöb görbe alatt vannak. Ez az a megfigyelés, ami a pszichoakusztikai

tömörítés alapja - ami lehetővé teszi, hogy több zajt engedélyezzünk a hangokban az erős elfedési hangok közelében.

A trükk természetesen az, hogy találjunk egy olyan algoritmust, ami megbízhatóan elárulja a tömörítőnek, hogy mely hangokat lehet átvinni több zajjal. Általában ezek a tömörítők a mintavételezett jelek frekvenciatartományba való transzformálásával dolgoznak, aztán alkalmazzák a pszichoakusztikát, minden frekvencia sávot kvantálnak és elküldenek. A dekódolónak csak a továbbított frekvencia sávokat kell visszaalakítania az időtartományba.

Két fő szabvány van, ami ezen elvek felhasználásával működik - a Dolby AC és az MPEG Audio. A két alkalmazás komplexitása és hatékonysága eltérő. A legfejlettebb alkalmazások a Dolby AC-3 és az MPEG1 Layer III, amelyek nagyon hasonlítanak egymásra, de míg az AC-3 optimalizálva van filmek hangsávjának többcsatornás kódolására, a normál sztereó jeleknél az MPEG-1 a hatékonyabb.

7.2. Az MPEG/audio tömörítés

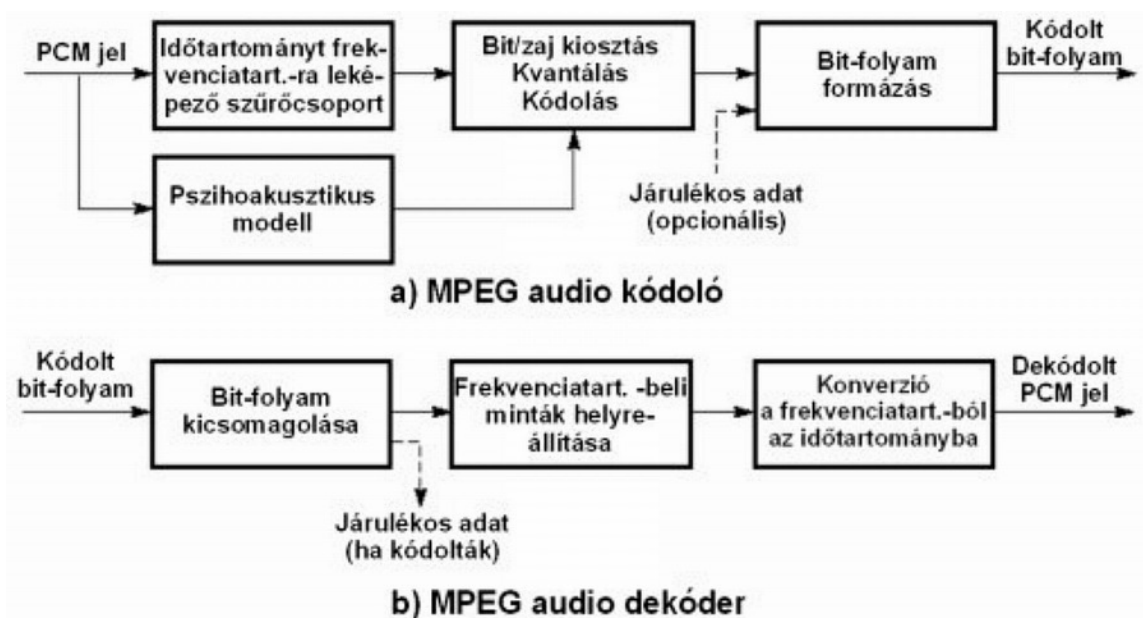
Az MPEG/audio egy általános audio tömörítési szabvány. Azon kódoló modellekkel ellentétben, amelyek speciálisan beszédjelekre vannak kialakítva, az MPEG/audio kódoló a tömörítést úgy végzi, hogy nem feltételez semmit az audio forrás jellegéről. A kódoló inkább az emberi hallórendszer észlelhető korlátait hasznosítja. A tömörítés főként az audio jel érzékelhetően irreleváns részeinek eltávolításával történik. Ezen részek eltávolítása csak olyan torzításokat eredményez, amelyeket nem észlelhetünk, így az MPEG/audio bármely az emberi fül által hallható jelet visszaadja. Mivel ez egy általános tömörítő, többféle tömörítési módot tesz lehetővé:

- a mintavételezési sebesség lehet 32, 44.1 vagy 48 kHz,
- a tömörített bitfolyam egy vagy két audio csatornára támaszkodhat a négy lehetséges módon:
 - monophonic mód egyetlen audio csatorna esetén,
 - dual-monophonic mód két független audio csatorna esetén (ez nem azonos a stereo móddal),
 - stereo mód sztereó csatornák esetén csatornák közötti bitmegosztással,
 - joint-stereo mód, ami vagy kihasználja a sztereó csatornák közötti korrelációt, vagy a csatornák közötti fáziskülönbség irrelevanciáját, vagy mindkettőt.
- a tömörített bitfolyam bitsebessége a néhány előredefiniált fix bitsebesség egyike lehet, amelynek tartománya 32 - 224 kbps per csatorna. A mintavételi sebességtől függően, ez tömörítési együtthatóként értelmezve 2.7 - 24. Ráadásul a szabvány gondoskodik egy „független” bitsebesség módról, amely az előre definiált bitsebességektől eltérő fix bitsebességet tesz lehetővé,
- az MPEG/audio három független tömörítési réteget kínál. A különbség a rétegek között a kodek komplexitása és a tömörített hang minősége:

- a Layer I a legegyszerűbb és a legalkalmasabb 128 kbps bitsebesség esetén,
 - a Layer II komplexitása közepes és a 128 kbps körüli bitsebességeket célozza,
 - a Layer III a legösszetettebb, de ez nyújtja a legjobb audio minőséget, főként 64 kbps körüli bitsebességeknél. Ez a réteg különösen alkalmas ISDN-en keresztül történő audio átvitelekre.
- a kódolt bitfolyam tartalmaz egy opcionális ciklikus redundancia ellenőrző (CRC) hibadetektáló kódot,
 - az MPEG/audio gondoskodik a bitfolyam belsejében lévő járulékos adatok tárolásának lehetőségéről.

Az MPEG/audio tömörítés kezdete a kvantálás. Bár a kvantálás veszteséges, ez az algoritmus veszteségmentesként érzékelt tömörítést ad. Az MPEG/audio bizottság kiterjedt, szubjektív vizsgálatokat végzett a szabvány kifejlesztése során. A tesztek azt mutatták, hogy még a 6:1 tömörítési aránynál (sztereó, 16 bit/minta, 48 kHz-es mintavételezési sebesség, 256 kbps-os tömörítés), és optimális tömörítési feltételek mellett, a szakértelm-mel rendelkező vizsgált személyek sem voltak képesek statisztikailag jelentős különbséget tenni a kódolt és az eredeti audio között. Továbbá ezeket a hangokat speciálisan választották, hogy nehéz legyen őket tömöríteni.

Az MPEG-1 audio kódoló és dekódoló:



Az időtartományt frekvenciartományra leképező szűrőcsoport egy többsávú szűrő, ami a jelet 32 részsávra képezi le. A részsávok egyenlő szélességűek, tehát nem azonosak a kritikus sávokkal. A szomszédos szűrők jelentős frekvenciaátlapolással rendelkeznek. Egy adott frekvenciájú jel így két szomszédos szűrőt is befolyásolhat.



Az MPEG/audio algoritmus úgy tömöríti az audio adatokat, hogy az audio jel akusztikailag irreleváns részeit eltávolítja. Kihasználja az emberi hallórendszer azon tulajdonságát, hogy az képtelen meghallani a kvantálási zajt, ha az el van fedve.

A gyakorlati eredmények is azt mutatják, hogy az emberi hallórendszer korlátozott, függ a frekvenciától, a hangot frekvenciától függő komponensekre bontja. A frekvenciafüggőség felbontásának finomsága a kritikus sáv szélességek segítségével modellezhető. Az emberi hallórendszer kiegyenlíti az egy kritikus sávon belül lévő változó hangjelkomponenseket.

Az emberi hallórendszer frekvenciafüggő felbontási képessége miatt a zaj elfedési küszöb bármely frekvencián kizárólag az adott frekvencia korlátozott sáv szélességű szomszédságán belül lévő jel energiájától függ. Az MPEG/audio úgy dolgozik, hogy felosztja az audio jelet frekvencia részsávokra, aztán minden részsávot kvantál az adott sávon belüli kvantálási zaj hallhatóságának megfelelően. A leghatékonyabb tömörítés esetén minden sávot úgy kell kvantálni, hogy ne használjunk annál több szintet, mint amennyi a kvantálási zaj hallhatatlanná tételéhez szükséges.

Pszichoakusztikai modell. A pszichoakusztikai modell elemzi az audio jelet, és kiszámolja a frekvencia függvényében előálló zajelfedési mennyiséget. Egy adott jel komponens elfedése függ a frekvenciájától és a hangosságától. A kódoló felhasználja ezt az információt, hogy eldöntse, hogyan tudja a legjobban reprezentálni a bemenő audio jelet a kódoló bitek korlátozott számával. Az MPEG/audio sztenderdben két lehetséges pszichoakusztikai modell implementáció található. Az első kevésbé összetett, mint a második, és több kompromisszumot köt a számítások egyszerűsítése érdekében. Bármelyik modell használható a tömörítés bármelyik rétege esetén, azonban csak a második modell tartalmaz olyan speciális komponenseket, amelyek segítségével a Layer III alkalmazható.

A pszichoakusztikai modell implementációja jelentős szabadsággal rendelkezik. A modell pontossága függ a megcélzott tömörítési együtthatótól és a tervezett alkalmazástól. A tömörítés alacsony szintjeinél, ahol bőséges készlet van kódbitekből, a pszichoakusztikai modell alkalmazását teljes mértékben mellőzhetjük általános használat esetén. Ekkor a bit allokációs folyamat iteratíván kiosztja a biteket a legalacsonyabb jel-zaj aránnyal rendelkező részsávokhoz. A zene archiválása során a pszichoakusztikai modellt sokkal szigorúbban kell alkalmazni.

Itt látható valamennyi réteg esetén a pszichoakusztikai számítások alaplépéseinek általános áttekintése. A két modell közötti különbségeket hangsúlyozni fogjuk:

- Audio adatok időbeli kiegyenlítése. Keretenként egy pszichoakusztikai kiértékelés van. A pszichoakusztikai modellnek küldött audio adatnak és a kódolással kapott audio adatnak időben meg kell egyeznie. A szűrés viszont időbeli késleltetést eredményez. A pszichoakusztikai modellnek kezelnie kell az audio adatoknak a szűrő bankon keresztül történő átvitelével kapott késleltetést és gondoskodnia kell egy adateltolásról, így a fontos adatok a pszichoakusztikai elemző ablakban fognak összpontosulni. Például ha az első pszichoakusztikai modellt használjuk Layer I esetén, a szűrő bankon keresztül történő átvitel késleltetése 256 minta, és az 512 mintából álló elemző ablakban lévő Layer I keret 384 mintája köré kell az eltolásnak összpontosulnia $(512-384)/2=64$ pont. A nettó eltolás 320 pont, ennyivel kell kiegyenlíteni a pszichoakusztikai modell adatokat, hogy időben egybeessenek a szűrő bank outputokkal.
- Audio konvertálása frekvencia tartományra. A pszichoakusztikai modellnek egy különálló, független, idő-frekvencia beosztást kellene használnia a többfázisú szűrő bank helyett, mert az elfedési küszöbök pontos kiszámításához finomabb frekvencia felbontásra van szükség. Mindkét pszichoakusztikai modell Fourier-transzformáltat használ ehhez a felbontáshoz. Egy szűrt Hann-súlyozás, amit az audio adatra a Fourier-transzformáció előtt alkalmazunk, az adatot úgy szabályozza, hogy a transzformáló ablak szegélyének hatásait csökkentsük.

Az első pszichoakusztikai modell egy 512 mintából álló elemző ablakot használ a Layer I-nél, és egy 1024 mintából álló ablakot a Layer II-nél és III-nál. Mivel egy Layer I keretben csak 384 minta van, az 512 mintából álló ablak elegendő terjedelmet biztosít. Itt a kisebb ablakméret csökkenti a számítások számát. A Layer II és III 1152 mintából álló keretet használ, így az 1024 mintából álló ablak nem fedi ezt le teljesen. Az elemző ablakon kívülre eső mintáknak általában nem lesz hatásuk a pszichoakusztikai kiértékelésre.

A második modell 1024 mintából álló ablakot használ mindegyik réteg esetén. A Layer I-nél a modell a keret 384 mintája köré összpontosul a pszichoakusztikai ablakban, ahogy korábban is tárgyaltuk. Layer II és III-nál a modell két 1024 pontú pszichoakusztikai kalkulációt végez minden egyes keretre. Az első számítás az elemző ablakban lévő 1152 minta első felére, a második számítás a második felére összpontosul. A modell kombinálja a két számítás eredményét úgy, minden részsávban a két jel-zaj arány közül a magasabbat használja. Ez a folyamat valójában a két zajelfedési küszöb közül az alacsonyabbat választja ki mindegyik részsáv esetén.

- A csoportosított spektrum értékek feldolgozása a kritikus sáv szélességek vonatkozásában. A pszichoakusztikai számítások egyszerűsítése érdekében mindkét modell feldolgozza az érzékelhető mennyiségű frekvenciaértékeket.
- Spektrum értékek elkülönítése tonális és nem tonális (zaj) komponensekre. Mindkét modell azonosítja és elkülöníti az audio jel tonális és nem tonális komponenseit, mert a jel ezen két típusának elfedési képességei különböznek.

Az első pszichoakusztikai modell az audio teljesítmény spektrumának lokális csúcsai alapján azonosítja a tonális komponenseket. Miután feldolgozta a tonális komponenseket, az első modell összeadja a maradék spektrum értékeit egy különálló nem tonális komponensbe kritikus sávonként. Ezen koncentrált nem tonális komponensek mindegyikének frekvencia indexe a határoló kritikus sáv geometriai átlagához eső legközelebbi érték.

A második pszichoakusztikai modell valójában sose különíti el a tonális és nem tonális komponenseket. Helyette kiszámít egy tonalitás indexet a frekvencia függvényeként. Ez az index méri mind a hang-szerű, mind a zaj-szerű komponenseket. Ezt az indexet a modell arra használja, hogy beszűrje a tiszta hangot-elfedő-zaj és a zajt-elfedő-tiszta hang értékek közé. A tonalitás index a jósolhatóságon alapul. A második modell az aktuális elemző ablakot megelőző két elemző ablakból használ fel adatokat, hogy megjósolja lineáris extrapolációval az aktuális ablaknál lévő komponens értékeit. A tonális komponensek jobban jósolhatóak, és ezért lesz magasabb tonalitás indexünk. Mivel ennek az eljárásnak sok adatra van szüksége, ezért még megfelelőbb, ha jobban elkülönítjük a tonális és nem tonális komponenseket, mint azt az első modell eljárása tette.

- Kiterjesztési függvény alkalmazása. Egy adott jel elfedési képessége kiterjed a környező kritikus sávokra is. A modell megállapítja a zajelfedési küszöböket egy empirikusan meghatározott elfedés felhasználásával (első modell), vagy egy kiterjesztési függvény felhasználásával (második modell).
- Egy alacsonyabb határ beállítása a küszöb értékek esetén. Mindkét modell tartalmaz egy empirikusan meghatározott abszolút elfedési küszöböt, az abszolút hallásküszöböt. Ez a küszöb a hang hallhatóságának legalacsonyabb határa.
- Elfedési küszöb megkeresése minden részsávnál. Mindkét pszichoakusztikai modell magasabb frekvencia felbontással számolja ki az elfedési küszöböt, mint amelyet a többfázisú szűrő bank szolgáltat. Mindkét modellnek a részsáv küszöb értéket egy olyan értékből kell származtatnia, amely az adott részsávban lévő frekvenciáknál kiszámított elfedési küszöb sokaság egy lehetséges értéke.

Az első modell minden részsávban kiválasztja a minimális elfedési küszöböt. Míg ez a közelítés megfelel az alacsonyabb frekvenciájú részsávoknál, ahol a részsáv keskeny a kritikus sávhoz viszonyítva, pontatlan lehet magasabb frekvenciájú részsávoknál, mert a kritikus sávok és a részsávok megfeleltetése itt más. Ezek a pontatlanságok azért merülnek fel, mert az első modell a nem tonális komponensek mindegyikére koncentrált, minden kritikus sávban külön értékben tárolja őket egy külön frekvencián. Az első modell valójában a nem tonális komponenseket egy tonális komponensformába transzformálja. Ez a közelítés egy újabb kompromisszum a számítások csökkentése érdekében.

A második modell csak azon sávok esetén választja ki az elfedési küszöbök minimumát, ahol a részsáv széles a kritikus sávhoz képest az adott frekvencia tartományban. Az elfedési küszöbök átlagát használja ott, ahol a sáv keskeny a kritikus sávhoz

képest. A második modell pontosabb a magasabb frekvenciájú részsávok esetén, mert nem koncentrálna a nem tonális komponensekre.

- A jel-zaj arány kiszámítása. A pszichoakusztikai modell kiszámítja a jel-zaj arányt a részsávban (vagy Layer III esetén a sávcsoporthoz) lévő jel energia és a részsávban lévő minimális elfedési küszöb arányaként. A modell ezt az értéket adja át a kódoló bit (vagy zaj) allokációs szekciónak.

A folyamat következő lépései a kvantálás, a kódolás és a bitfolyam formázása. Mivel minket csak a pszichoakusztika érdekel, ezért ezekkel a lépésekkel nem foglalkozunk részletesebben.

A legtöbb audio konstans bitsebességgel (CBR) van tömörítve, ami azt jelenti, hogy minden másodpercben ugyanaz a bitmennyiség áll rendelkezésre. Azonban nyilvánvaló, hogy az audio jel minden, csak nem konstans. Vannak csendes és hangos, bonyolult és egyszerű szakaszok, de végül, ha konstans bitsebességgel kódolunk, akkor ezeket mindet ugyanazzal a bitsebességgel írjuk le.

Ezért a Xing Technology tervezett egy változó bitsebességű (VBR) kódoló rendszert, amiben minden keret bitsebessége skálázva van azon elv alapján, hogy némelyik résznek kevesebb, más részeknek több bitre van szüksége. Ez azt jelenti, hogy ugyanazon fájl méret esetén jobb minőségű audiot kaphatunk.

7.3. Speex és Vorbis

A Speex a Xiph.org által kifejlesztett veszteséges hangtömörítési eljárás. Ez az eljárás a beszéd tömörítésére lett kifejlesztve, és mivel a Vorbis pszichoakusztikai modelljét használja fel, azt fogjuk most bemutatni.

Adott egy 44.1 kHz-es sztereó input (amely a sztenderd CD audió mintavételi frekvencia). A kódoló ebből 45-től 500 kbps-ig állít elő kimenetet a minőségi beállításoktól függően. Ez azonban csak közelítés, mivel a Vorbis változó bitsebességgel (VBR) dolgozik. A bitsebesség így mintáról mintára változhat.

A Vorbis felhasználja a módosított diszkrét koszinusz transzformációt (MDCT) a hangok időtartományból frekvenciatartományba való átkonvertálásánál. A kapott frekvenciatartomány felosztható zajszintre és maradék komponensekre.

A Vorbis pszichoakusztikai modellje úgy működik, hogy egy körülbelül egy Bark méretű mozgó ablakkal kisimítja a logaritmikus spektrumot, és ezáltal konstruál egy geometriai középértéket és egy burkoló kísért. A két görbe közötti különbség tonalitási becslést eredményez a spektrum adott sávjánál. A burkoló görbét tömöríti és kiterjeszti a távolságnak megfelelően, hogy a spektrális súlyozásnál használt zajelfedési görbét közvetlenül kiszámíthassa.

7.3.1. Pszichoakusztikai alapötletek

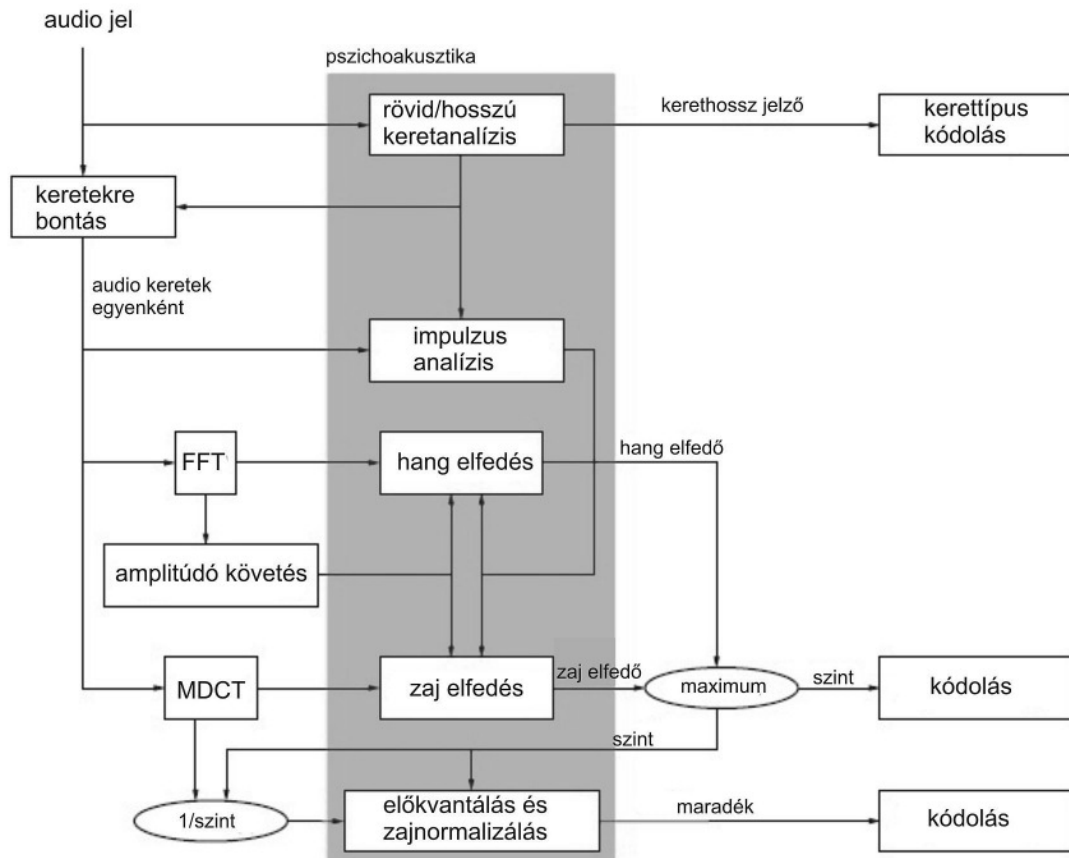
Az audió kódolások kétféle pszichoakusztikai módszert használnak. Az első egy szeparáló mechanizmus, ami a speciális audió tulajdonságok relatív jelentőségét kvantálja úgy, hogy

a kódolás a legfontosabb reprezentálás irányába lesz súlyozva. Ez az analitikus optimalizáció a kódolási oldalon jelenik meg. A legtöbb kodek, beleértve a Vorbis-t és a Speex-et is, ebbe a kategóriába tartozik.

Másodszor, minden audió kodek támaszkodik néhány pszichoakusztikai mennyiségre, amelyet a kodek specifikációjába építenek be. Az ilyen bedrótolt pszichoakusztika mind a kódoló, mind a dekódoló mechanizmusainak része lesz. Néhány jelenlegi kodek teljesen kihagyja a korábbi opcionális pszichoakusztikát, teljes mértékben a kodek bedrótolt pszichoakusztikájára támaszkodik. Ezekben a kódokban a pszichoakusztikai modellt nem lehet megváltoztatni a bitfolyam megváltoztatása nélkül.

7.3.2. A Vorbis modellje

A Vorbis kodek a kódolás oldali pszichoakusztikai heurisztikákra támaszkodik, az eredményes bitfolyamok előállításában.



7.1. ábra. A Vorbis kódoló blokkdiagrammja mono input adatfolyamra alkalmazva. A sötét rész tartalmazza a pszichoakusztika elemeit

Minden keret audió spektrumaiból kiszámít egy szint görbét, amelyet egység-felbontási fehérítő szűrőként használunk. A fehérítés a jelforrást úgy transzformálja, hogy az a fehér

zajhoz közelebb jelenjen meg. A szintet a spektrum normalizálására használjuk, így a normalizált spektrum közvetlen lineáris kvantálása érzékelhetően megfelelő. Így a szintnek hasonló funkciója van, mint a Speex-ben és más CELP kodekben használt egyszerűbben megfigyelhető súlyozó filternek. Azonban a Vorbis esetében a szint görbe a dekódolásnál is használható, míg a CELP a súlyozó filtert csak a kódkönyvben lévő legjobb bejegyzés keresésének optimalizálásánál használja.

A Vorbis kódoló által használt keretről keretre alkalmazott pszichoakusztikai metrikák, amelyek kiszámítják a szintet, négy fő kategóriára bonthatók: hang elfedés, zaj elfedés, zajnormalizálás és impulzus analízis.

Hang elfedés. A Vorbis kódoló referenciában szereplő hang elfedés egy egyszerű hang-hang elfedési motor, amely egyetlen spektrális görbét állít elő keretenként. Ez a görbe reprezentálja az érzékelés küszöbét az aktuális keretben lévő tiszta hangnál. A görbét tonális elfedési görbék szuperpozíciójából kapjuk.

A hang-hang elfedési analízis és a tonális elfedési görbéknek nincs igazán jelentőségük a modern Vorbis kodekben.

Zaj elfedés. A Vorbisban szereplő zaj elfedés egy rosszul elnevezett dolog, annak ellenére, hogy egy egyszerű zaj-zaj elfedésként kezdődik a fenti hang-hang elfedéshez hasonlóan, idővel egyre inkább foglalkozik egy olyan görbe konstruálásával, ami a spektrumban lévő zajenergia burkolóját közelíti. Ezt a görbét a geometriai középértékből és egy körülbelül egy Bark méretű mozgó ablakkal kisimított a logaritmikus spektrum által előállított burkoló kísérből számítjuk ki. A burkoló görbét tömörítjük és kiterjesztjük a burkoló és a középérték közötti távolságnak megfelelően. A nagyobb távolságok nagyobb tonálisra, a kisebb távolságok nagyobb zajosságra utalnak. Ennek megfelelően ezeknek a görbéknek az amplitúdója a nagyobb tonálisok területén kisebb. A Vorbis ezután hozzáad egy bedrótozott különbségi görbét (a zaj offszetet) a tömörített és kiterjesztett burkolóhoz, így végül előáll a zaj elfedés görbe.

Zajnormalizálás. A zaj normalizálás az elemzés része, és a zaj elfedésnél leírt zaj energiát kezeli, melyet külön lépésben végzünk. A normalizáció célja az, hogy megőrizze a megközelítőleg szélessávú energiát a kvantálás folyamán, főleg nagyon alacsony jel-zaj arányoknál. Olyan mechanizmusnak tekinthető, ami megőrzi a zaj elfedés célját a kvantálási eljárás alatt.

Impulzus analízis. Az impulzus analízis különböző metrikákra hivatkozik, amelyek egy audió keretben olyan időben erősen korlátozott eseményeket jellemeznek, mint a hirtelen impulzus vagy lökeshullám a hangban. A Vorbis kódoló korábbi verzióiban lökeshullámokat analizál, és előrejelzi a lehetőségeket, hogy meghatározzuk, mikor használjunk rövid és hosszú kereteket, felhasználva ezeket a metrikákat a keret belső kódoló algoritmusában. Az audió ezen fajtájának legnyilvánvalóbb példája a beszédhang.

Mint a zaj normalizációnál, az impulzuselemzést is egy kodek átalakítás fajta teszt szükségessé.

Irodalomjegyzék

- [1] E. Zwicker, H. Fastl: *Psychoacoustics : facts and models*.
Második, javított kiadás. Springer, 1999.
- [2] Tarnóczy T.: *Zenei akusztika*.
Zeneműkiadó, 1982.
- [3] Budó Á.: *Kísérleti fizika : 1. kötet. Mechanika, hangtan, hőtan*.
Tizedik kiadás. Tankönyvkiadó, 1989.
- [4] Gordos G., Takács Gy.: *Digitális beszédfeldolgozás*.
Műszaki Könyvkiadó, 1983.
- [5] *Pszichoakusztika*.
<http://www.gdf-ri.hu/targy/mikroszg/Diploma/html/05pszihoakusztika.htm>
- [6] Lois L.: *Audió alapok*.
<http://www.hit.bme.hu/musor/egyeb/audio-al.pdf>
- [7] *Beszédinformációs rendszerek - órajegyzet - 2005*.
http://speechlab.tmit.bme.hu/download/Beszinfo_1_fejezet_2005.pdf,
2005.
- [8] Osvay K.: *Előadás 7*
[http://titan.physx.u-szeged.hu/~opthome/optics/oktatas/
Zenei_akusztika/Eloadas7.pdf](http://titan.physx.u-szeged.hu/~opthome/optics/oktatas/Zenei_akusztika/Eloadas7.pdf), 2005.
- [9] Szappanos A.: *Hallás*.
<http://www.kfteam.hu/iromanyok/hallas.pdf>, 2002.
- [10] Szabó G.: *Analóg jelek digitális feldolgozása és továbbítása*.
http://www1.kka.bme.hu/irtech/public_html/I1B/adda_ea.pdf, 2001.
- [11] B. Wesén: *Compression of digital audio*.
<http://www.sparta.lu.se/~bjorn/whitney/compress.htm#Psychoacoustics>,
1997.
- [12] *MPEG/audio tömörítés*.
<http://www.gdf-ri.hu/targy/mikroszg/Diploma/html/05MPEGkod.htm>

- [13] *The Ins and Outs of Audio Compression.*
<http://www.animemusicvideos.org/guides/avtech/audio2.html>, 2002.
- [14] D. Pan: *A Tutorial on MPEG/Audio Compression.*
<http://www.cs.columbia.edu/~coms6181/slides/6R/mpegaud.pdf>, 1996.
- [15] J.-M. Valin: *Improved Noise Weighting in CELP Coding of Speech - Applying the Vorbis Psychoacoustic Model To Speex.*
http://people.xiph.org/~jm/papers/aes120_speex_vorbis.pdf, 2006.