

Vnímanie zvuku, psychoakustika, audio a rečové signály

Juraj Kačur

Obsah

- Audio a rečové signály
- Vlastnosti / delenia
- Ľudský sluchový systém
- Psychoakustika

2

Vzhľadom na to že naša komunikácia prirodzene prebieha pomocou ľudských zmyslov je treba pre korektný návrh rozhrania človek stroj postupne skúmať každý zmysel a to zo všetkých relevantných aspektov. V tejto stati sa preto budeme venovať komunikácii pomocou zvukových signálov. Ďalej budú preto postupne rozoberane vlastnosti audio signálov, ich delenia, ľudsky sluchový systém, využitie audio signálov ako aj psychoakustika ktorá má pre ľudí veľký význam.

Audio signály

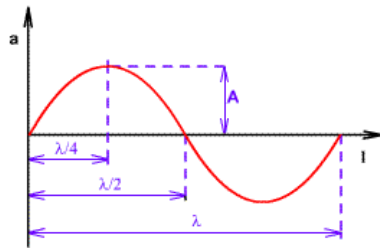
- Mechanické vlnenie pružného prostredia
 - Najčastejšie vzduch
- Pozdĺžna zmena tlaku v smere šírenia
 - Nedochádza k pohybu prostredia
- Vyjadrenie harmonického signálu v 1D prostredí
 - Riešenie vlnovej rovnice v tvare

$$a(t, x) = A(x) \cos\left(\omega t + \varphi - \frac{2\pi}{\lambda} x\right) = A(x) \cos\left(\omega t + \varphi - \frac{2\pi}{\lambda} ct\right)$$

3

Zvukové signály sa primárne šíria vzduchom i keď ich prenos je možný aj v iných prostrediach (latkách), napr. voda, kov, sklo. Tieto signály sú reprezentované mechanickým vlnením danej latky. V nasledovnom texte budú spomenuté základné vlastnosti takéhoto vlnenia ako aj jeho matematická reprezentácia.

Audio signály



C- rýchlosť šírenia vlny v prostredí

λ - vlnová dĺžka

T -perióda vlnenia

ω - kruhová frekvencia

Platí:

$$\lambda = cT$$
$$\omega = 2\pi/T$$

4

Grafické zobrazenia vlny (sínus) a vysvetlenie vlnovej dĺžky, frekvencie, periódy a rýchlosti šírenia vlny v danom prostredí.

Audio signály

- Zvuk sa šíri všetkými smermi rovnako
- Na prekážke dochádza k
 - Odrazu
 - Pohltenu
 - Ohýbaniu
- Odraz a pohltenie zvuku súvisí s materiálom
- Pri pohltení sa mení v látke na teplo

5

Jedny so základných a známych vlastností zvuku sú nasledovne.

Audio signály

- Rýchlosť zvuku súvisí s typom prostredia
 - Závisí od teploty
 - Nezávisí od tlaku vzduchu ani frekvencie vlnenia

Prostredie	Rýchlosť (m.s ⁻¹)
Suchý vzduch (0 °C)	331,4
Suchý vzduch (25 °C)	346,3
Destilovaná voda (25 °C)	1 497
Morská voda (13 °C)	1 500
Ľad (13 °C)	3 200
Med' (20 °C)	4 720
Sklo (20 °C)	5 200
Hliník (20 °C)	6 400

6

Rýchlosť šírenia zvuku je značne závislá od materiálov a aj teploty; tu sú pre zaujímavosť zobrazené niektoré z nich

Audio signály

- Hlasitosť zvuku súvisí s veľkosťou zmeny tlaku v danom mieste
- Objektívne miery „intenzity, sily“ zvuku
 - Dajú sa zmerať pomocou základných jednotiek
- Subjektívne miery „intenzity, sily“ zvuku
 - Súvisia s tým ako človek vníma „intenzitu“-hlasitosť
 - Nie je lineárna v závislosti od fyzicky nameraných parametrov
 - Závisia od frekvencie
 - Líšia sa od jednotlivca k jednotlivcovi (priemerný) ₇

Kvôli bližšiemu skúmaniu a popisu charakteristík zvukových signalov bolo potrebné definovať niektoré miery. Tie rozlišujeme na objektívne a subjektívne. Toto delenie je potrebné vzhľadom na to, že i keď objektívne miery sa dajú presne a jednoznačne zmerať (s danou presnosťou merania) ľudské ucho spolu s mozgom ich môžu interpretovať odlišne. V nasledovnom sú niektoré z týchto mier spomenuté bližšie najmä tie týkajúce sa vnemu hlasitosti.

Audio signály

- Akustický tlak, zmena tlaku spôsobená zvukom, jednotka Pascal
 - Má veľký rozsah a preto na jeho rozumne vyjadrenie je potrebná nelineárna mierka
- Hladina akustického tlaku (sound pressure level, SPL)
 - Vyjadruje sa v dB, komprimuje rozsah hodnôt do rozumných limitov

$$Spl = 20 \log \left(\frac{p_1}{p_0} \right)$$

Audio signály

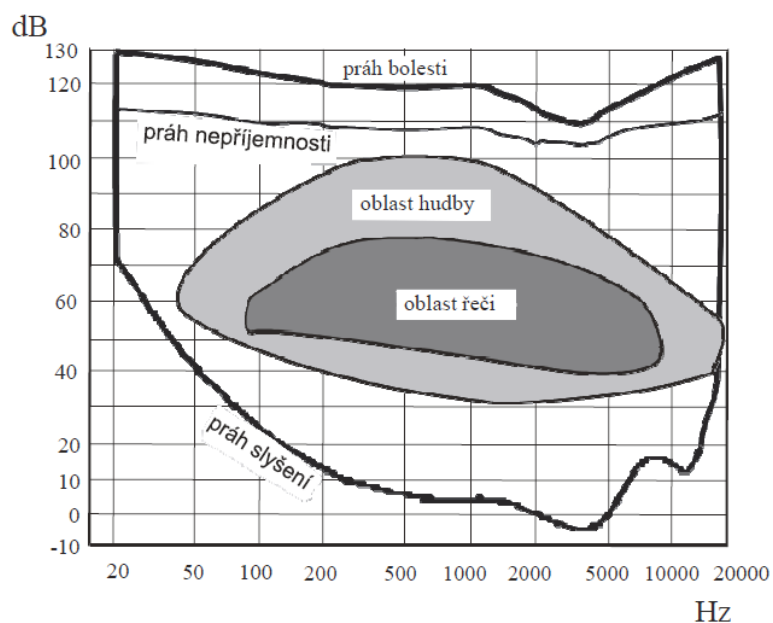
- p_0 referenčný tlak
 - prah počuteľnosti
 - 1kHz tón
 - $p_0 = 20,4 \mu\text{Pa}$ (zvolený za štandard)
- p_1 nameraný akustický tlak
- Je to pomerová jednotka vzťahnutá k referenčnému tlaku
- Intenzita zvuku
 - množstvo zvukovej energie, ktorá prejde danou plochou za jednotku času (W/m^2)
 - klesá so štvorcom vzdialenosti

Audio signály

L_p [dB(SPL)]	p [Pa]	zdroj zvuku, udalosť s danou hlasitosťou
0	0,000 020	<i>prah počuteľnosti</i> (zdravého jedinca, 1 kHz)
10	0,000 063	ľudský dych zo vzdialenosti 3 metrov
30	0,000 632	ticho v divadle, noc na púšti
40	0,002 000	ticho na sídlisku v noci, ľudský šepot
50	0,006 325	tichá kaviareň
60	0,020	zvyčajná hlasitosť normálnej konverzácie
80	0,200	vysávač zo vzdialenosti 1 meter
90	0,632	nákladné auto zo vzdialenosti 1 meter
100	2	diskotéka, televízor alebo veža pri najvyššej hlasitosti
120	20	rockový koncert, vzlet prúdového lietadla (100 m) telo cíti vibrácie pri nízkych frekvenciách
130	63	vojenský dychový orchester zblízka <i>prah bolesti pre všetky frekvencie</i>
150	632	prúdový motor (30 m) bolesť v hrudníku, rozmazané videnie <i>poškodzovanie sluchu pri všetkých frekvenciách</i>
180	20 000	motor rakety (30 m), dunenie vráskavca ozrutného (1 m) Výbuch sopky Krakatau (r. 1883) zo vzdialenosti 160 km! ¹⁰

V nasledovnej tabuľke sú vyjadrené úrovne tlaku [Pa], hladiny akustického tlaku [dB] a k nim zodpovedajúce zdroje zvukov o daných úrovniach hlasitosti, pre lepšie uvedomenie si daného problému. Z uvedených čísel je vidieť aké praktické je zavedenie nelineárnej pomerovej jednotky merania hlasitosti uvedenej v dB.

Audio signály

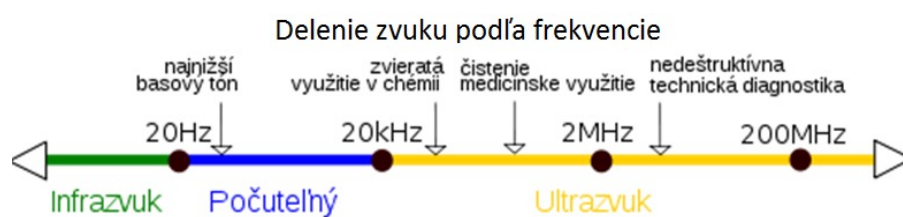


11

Nasledující obrázok ukazuje rozloženie typických zvukov ako reč a hudba v oblasti frekvencii a príslušných úrovni intenzity zvukov. Kvôli veľkému rozsahu frekvencii je aj frekvenčná mierka uvedená v logaritmickú mierke. Cella oblasť je ohraničená prahom počutia a bolesti čo sa týka rozsahov tlakov a maximálnym rozsahom frekvencií ktoré človek vie počuť.

Audio signály

- Delenie (rozsah)
 - Počuteľný zvuk : 20-20000 (Hz)
 - Infrazvuk : pod 20 (Hz)
 - Ultrazvuk: nad 20000(Hz)



12

Zvuk sa podľa frekvencií delí do 3 skupín, ktoré sú znázornené na obrázku.

Audio signály

- Vlastnosti a použitie
 - Počuteľný zvuk
 - Komunikácia, lokalizácia...
 - Infrazvuk
 - Ohyby za prekážkou
 - Dorozumievanie na dlhé vzdialenosti: veľryby v oceánoch
 - Má nepriaznivé účinky ak dlhodobo pôsobí
 - Ľudia ho môžu vnímať ako chvenie
 - Ultrazvuk
 - Ohýba a odráža sa na prechodoch pri rôznych materiálových
 - Detekcie nehomogenít prostredia, chyby materiálov
 - Meranie vzdialenosti, polohy
 - Čistenie predmetov- kavitácia

13

Každá skupina zvukov ma svoje špecifické vlastnosti a použitie, ktoré sú vysvetlene ďalej.

Audio signály

- Vznik a typy
 - Reč: hlasové orgány človeka
 - špecifické vlastnosti (300- 4000Hz)
 - Len určité postupnosti znelých a neznelých úsekov
 - Audio signály produkované hudobnými nástrojmi
 - Pre každý nástroj majú špecifické vlastnosti
 - Zvyčajne každý produkuje unikátnu množinu harmonických zložiek k základnému tónu
 - Harmonické zložky: násobky základnej periódy, určujú farbu zvuku
 - Každý ma aj iný nástup a dobeh (brnknutie do struny, úder klávesu, atď.)

14

Audio signály delíme aj podľa toho ako vznikli, resp. ich vlastnosti a použitia. Rozlišujeme nasledovne skupiny každá so svojimi významnými vlastnosťami.

Audio signály

- Vznik a typy
 - Zvuky prostredia (šumy):
 - Podľa typu prostredia existuje viacero typov šumov
 - Líšia sa frekvenčnými a časovo energetickými vlastnosťami
 - Známe typy:
 - Biele šumy
 - » majú všetky frekvencie rovnako zastúpene (kvantizačný šum, tepelný šum, atď.)
 - Ružový šum
 - » výkonové spektrum klesá s frekvenciou $P_x(f) = 1/f^a$, $a \approx 1$
 - » z hľadiska vnímania je prirodzenejší
 - » Má vyrovnanjšie výkonové rozloženie pre ľudský vnem
 - » Používa sa pre testovanie audio zariadení

15

Veľkú skupinu tvoria rôzne typy šumov, ktoré majú špecifické vlastnosti a použitie v praxi. Tie najznámejšie sú uvedené nižšie.

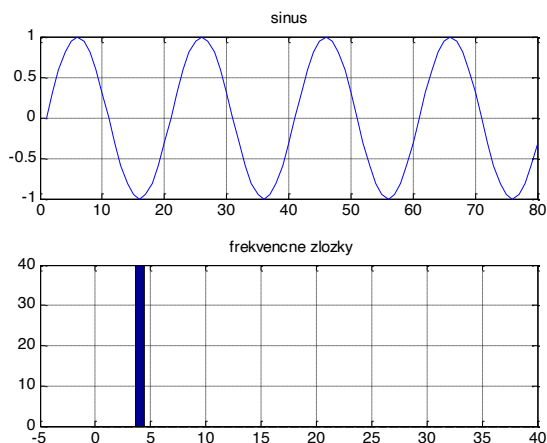
Audio signály

- Šedý šum
 - » Kopíruje psychoakustické vnímanie hlasitosti pre rôzne frekvencie
 - » 7-500Hz ploché spektrum, potom pokles o 3dB na oktávu
 - » „skutočný“ biely šum pre človeka
- Purpurový šum
 - » výkonové spektrum rastúce 6dB na oktávu
 - » Rastie proporcionálne s f^2
 - » Derivácia bieleho šumu
- Zelený šum
 - » dlhotrvajúcim priemerovaním hlukov pozadia v otvorenej krajine, „šum sveta“
 - » plochý od 500 až 2000Hz, potom klesá 9dB na oktávu. Od 500Hz nižšie klesá 6 dB na oktávu
 - » simulácia telekomunikačných kanálov

16

Audio signály

- Príklady harmonických zložiek periodických signálov s periódou 4
 - Sínus, harmonický signál, nemá iné harmonické zložky

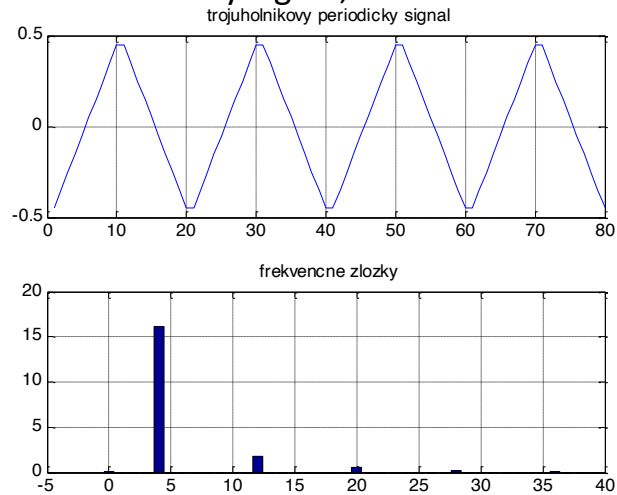


17

Na nasledovných obrázkoch sú uvedené 3 základné typy priebehov signálov (sinus, pila, oblzník) a k nim sú ukázané ich frekvenčné zložky (harmonické signály s danou frekvenciou) z ktorých daný signál skladá.

Audio signály

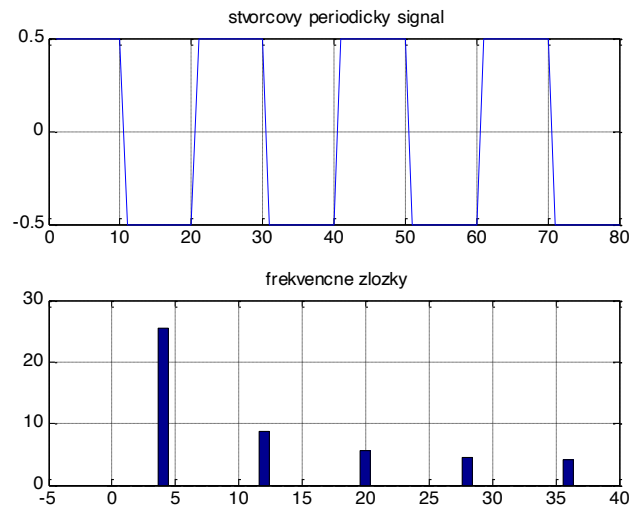
- Trojuholníkový signál, blízky ku sinus, ale nie je dokonale harmonický signál, má iné harmonické zložky



18

Audio signály

- Pravouhlý signál, odlišný od sinus, má aj iné harmonické zložky



19

Audio signály

- Skreslenie signálu

- Prítomnosť aditívneho šumu

- $y(t)=x(t)+sum(t)$
 - Miera zašumenia odstup signál šum (SNR)

$$SNR_x = 10 \log\left(\frac{P_{signal}}{P_{sum}}\right)$$

- Nelineárne skreslenie

- Vzniknú harmonické zložky (vyššie ako prvá)
 - Celkové harmonické skreslenie (ako sa líši od harmonického)

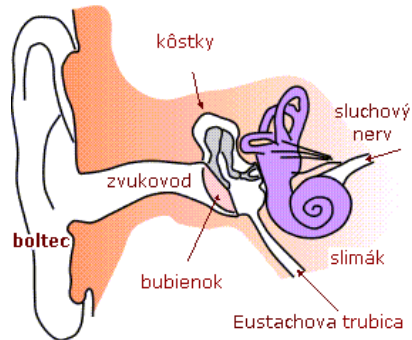
$$THD_x = \frac{\sum_{i=2} P_i}{P_1}$$

20

V reálnych prostrediach je užitočný signál degradovaný okolím. Existuje viacero degradácií, najčastejšie sa vyskytuje prítomnosť aditívneho sumu z okolia. Okrem toho môže nastať aj nelineárne skreslenie signálu (nahrávacie, reprodukčné zariadenia) kedy sa tvar signálu deformuje. Tým vznikajú tzv. vyššie harmonické zložky signálu.

Ľudský audio systém

- Konštrukcia ucha
- Ucho sa delí na 3 časti
 - Vonkajšie ucho
 - Boltec, ušnica- „zbiera“ zvuk
 - Vonkajší zvukovod, vedie zvuk
 - Stredné ucho
 - Ušný bubienok: mení zvuk na vibrácie
 - Kladivko, nákovka, strmienok: prenášajú vibrácie do stredného ucha



21

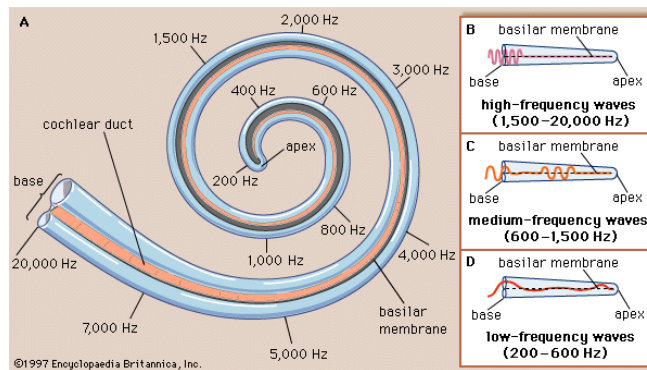
Na pochopenie ako spracováva a vníma človek zvuk je najprv potrebné objasnenie ako biologicko- mechanicky funguje zber zvukových podnetov z okolia, t.j. funkcia ucha. V ďalšom je preto stručne naznačený celý proces aj s príslušnými obrázkami, deleniami funkciami a časťami ucha.

Ľudský audio systém

– Vnútročné ucho

- Cochlea

- slimákovitá štruktúra obsahujúca tekutinu
- Pokrytá je vlásočnicovými bunkami
- Sú citlivé na pohyb, generujú nervové vzruchy
- Vzhľadom na tvar dochádza k filtrácii frekvencií na rôznych pozíciách



22

Jej tvar umožňuje mechanicky filtrovať zvuky podľa frekvencií. Na začiatku sú vysoké frekvencie a do hĺbky prechádzajú len nižšie frekvencie. Preto tým odkiaľ vzruchy prichádzajú vieme rozlíšiť frekvencie

Ľudský audio systém

- Vlásoknicové bunky generujú prirodzene vzruchy- pulzy
- Trvanie pulzu 0.2ms
- Prirodzená frekvencia generovania pulzov 50/s
- Počas akustickej stimulácie 150/s
- V prípade silnej až 1000/s
- Sluchový nerv
 - Vede vzruchy do mozgu
 - Nervy sú uchytene na povrch membrány
- Polkruhovitý kanálik
 - Poloha tela
 - priamo neslúži na vnem zvuku

Ľudský audio systém

- Popis činnosti
 - Zvuk sa nesie zvukovodom a rozhýbe ušný bubienok.
 - Ušný bubienok vibruje zvukom.
 - Zvukové vibrácie sa cez kostičky prenesú do kochley.
 - Zvukové vibrácie zapríčinia, že sa rozhýbe tekutina v kochley.

24

Samotná činnosť ľudského sluchového systému, t.j. ako je vstupný zvukový signál transformovaný na nervové vzruchy idúce do mozgu, je heslovite popísaná ďalej.

Ľudský audio systém

- Pohyb tekutiny ohýba vlásoknicové bunky. Tvorí nervový signál, ktorý zbiera sluchový nerv. Vlásoknicové bunky na jednom konci kochley posielajú nízkočrekvenčné zvuky a bunky na druhej strane kochley posielajú vysokočrekvenčné zvukové informácie.
- Sluchový nerv posielá signály do mozgu, ktorý ich interpretuje ako zvuk.

Psychoakustika

- Človek vníma odlišne rôzne vlastnosti, ktoré sú objektívne namerané
- Skúma fenomény ľudského zvukového vnemu a ich súvis s objektívnymi veličinami
 - Súvisí to s fyzickou konštrukciou audio systému
 - Mozgovou aktivitou
 - Je to dané evolúciou, lepšie prispôsobenie sa prostrediu a okolitým zvukom, potlačenie šumov, zvýraznenie podstatných dejov

26

Zistilo sa že je rozdiel medzi merateľnými vlastnosťami zvukových signálov a tým čo a ako ľudia vnímajú. Časť vedy ktorá sa tým zaoberá sa nazýva psychoakustika a je dôležitá na správnu konštrukciu zvukových/recovych senzorov.

Preto budú v ďalšom naznačené základne fenomény a poznatky o vnímaní zvuku ľuďmi.

Psychoakustika

- Vnímanie hlasitosti
 - Nelineárne vzhľadom na namerané veličiny (akustický tlak)
 - Nižšie intenzity sú zvýraznené, vyššie potlačené
 - Jednotky : fón, son
 - Vystupuje v nich 10-kový logaritmus pomeru výkonu
 - Napr. ak sa zvýši o 10 fónov, tak výkon je 10 násobný! (10w vs. 100w)

27

Existuje viacero jednotiek na subjektívneho ohodnotenia hlasitosti, napr. pre čisté tony alebo komplexne signály (viacej zložiek).

Psychoakustika

– Fón, [Ph]

- Súvisí s vnímaním intenzity čistých tónov
- Základ 1kHz
- $x \text{ [dB]} = x \text{ [Ph]}$

– Hlasitosť, [son]

- týka sa celkového vnímania hlasitosti
- 1 son pre 1kHz tón a 40dB slp

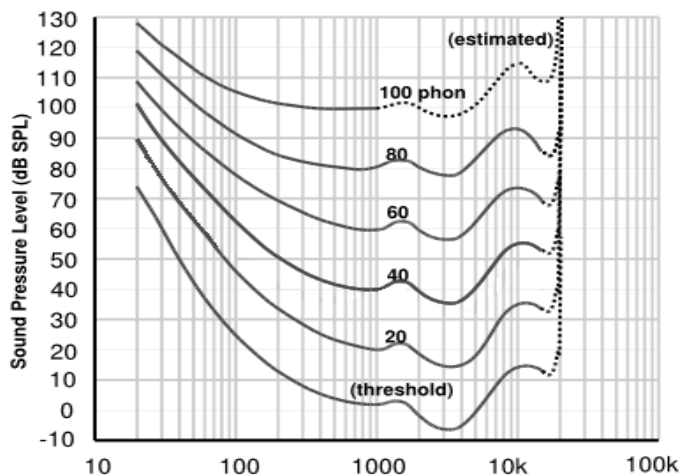
– Hlasitosť závisí od frekvencie

- Stanovila sa referenčná frekvencia 1kHz,
nastavovali sa výkony tak aby bola rovnaká
hlasitosť pre rôzne frekvencie

28

Psychoakustika

- Vnímanie hlasitosti závislé od frekvencie
- Krivky rovnakej hlasitosti pre rôzne úrovne SLP



Poznámky

- Krivky sa volajú izofóny
- Údaje sú získané priemerom na populácií a sú štandardizované
- Najvýraznejšie počujeme cca 2kHz
- Krivka sa mení od hlasitosti
- rozsah 20-20000, dospelí počujú do 16000

29

Ucho vníma hlasitosti rôzne a to v závislosti od frekvencie. Na obrázku sa naznačene krivky (izogóny) úrovne akustického tlaku, v závislosti od frekvencie. Pozdĺž každej izofony človek vníma hlasitosť rovnako. Z toho vyplýva že najcitlivejší sme na frekvencie okolo 2kHz. Tieto krivky sú individuálne ale sú ručne ako priemer z populácie.

Psychoakustika

Poznámky

- Krivky sa volajú izofóny
- Údaje sú získané priemerom na populácií a sú štandardizované
- Merané pre tóny
- Najvýraznejšie počujeme cca 2khz
- Krivky sa menia od hlasitosti
- Najnižšia krivka je **absolútny prah počuteľnosti**
- Najvyššia je **prah bolesti**
- rozsah 20-20000, dospelí počujú do 16000

30

Psychoakustické miery frekvencie

- Najmenší vnímateľný rozdiel vo výške tónov sa mení od výšky daného tónu
 - Napr. do 1kHz je rozdiel 2-3 Hz, pri 7kHz je to 35 Hz
- Vnímanie výšky tónu sa mení s frekvenciou
 - Merania: ľudia mali nastavovať vyššie a nižšie tóny
 - Referenčný bol 1kHz
 - Jednotlivé „subjektívne“ nastavenia sa priemerovali na populácií
 - Mel psychoakustická škála vzhľadom na vnímanie výšky tónu

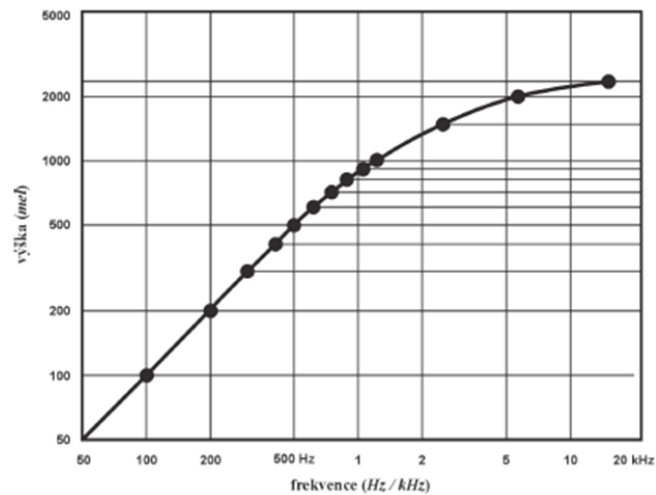
31

Okrem samotného rozdielného vnímania hlasitosti je rozdiel a vo vnímaní samotnej výšky frekvencie. Meraniami na populácii a priemerovaním sa zistilo že v skutočnosti s rastúcou frekvenciou od 1kHz človek vníma výšku frekvencie menej ako v skutočnosti je. Preto sa robili rôzne pokusy ako namapovať skutočnú frekvenciu na vnímanú a z nich vyšlo viacero frekvenčných skal- mier. Najznámejšia ktorá meria len jeden ton je Melova škála.

Psychoakustické miery frekvencie

- Vzťah medzi psychoakustickou škálou a Hz

$$f_{\text{Mel}} = 2595 * \log_{10}\left(1 + \frac{f}{700}\right)$$



32

Psychoakustické miery frekvencie

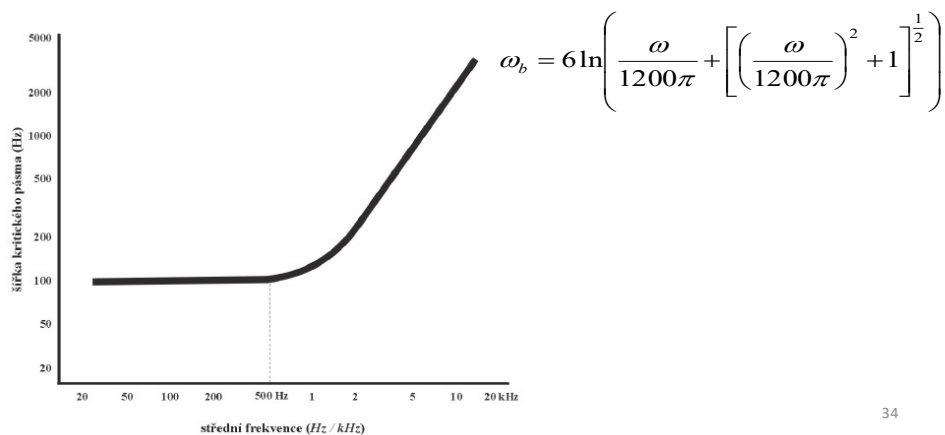
- Kritické šírky pásma
 - Kedy dva blízke tóny už vyvolajú iný vzruch na bazilárnej membráne (hlasitosť)
 - Kritické pásmo je max. rozsah frekvencií, ktoré sa ešte prejavia rovnako
 - Závislé od frekvencie
- Barkova škála
 - Okolo tónu bol vysielaný úzkopásmový šum
 - Postupne sa zvyšovala šírka šumu
 - Keď šum už bol hlasitosťou iný ako tón, to určilo šírku pásma
 - Bola stanovená na 1 bark

33

Ďalšie miery uvádzajú o existencii viacerých tónov súčasne a ich ovplyvňovaní sa navzájom. Sú založené na tzv. Kritických pásmach, ktoré ak sú prekročené už vnímame daný zvuk inak (hlasnejšie), resp. dochádza k maskovaniu prísušného tónu. Nevýznamnejšie sú Barkova škála a ERB škála, ktoré sú v stručnosti naznačené a opísané ďalej.

Psychoakustické miery frekvencie

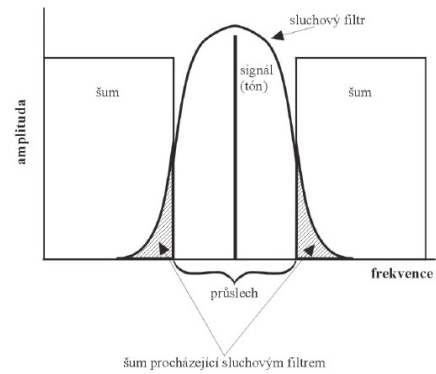
- Existuje viacero mat. aproximácií medzi Barkovou škálou a Hz
- Kritické šírky pásma pre jednotlivé frekvencie



34

Psychoakustické miery frekvencie

- ERB škála
 - Equal rectangle bandwidth
 - Tiež kritické pásma
 - K tónu sa pridával obojstranne šum
 - Keď začal byť tón maskovaný to stanovilo kritické pásmo



Frekvenčné maskovanie

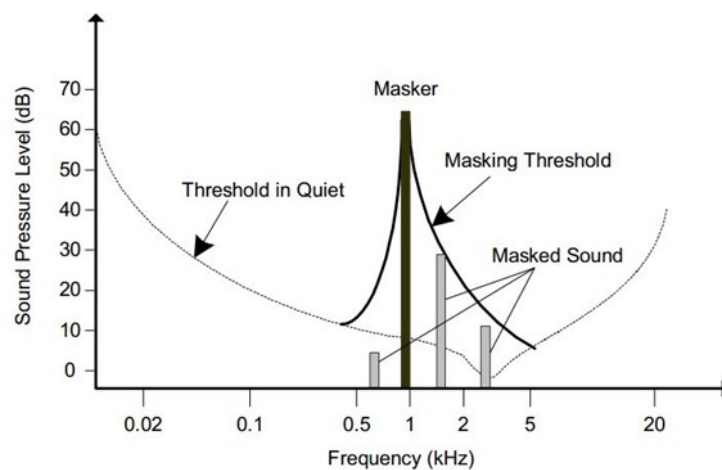
- Vnem jednotlivých tónov je ovplyvnený druhými
- Tým viac čím sú bližšie
- Väčší maskuje menšie tóny v okolí
- Je to komplexná funkcia závislá na frekvencií, veľkosti tónov a počte tónov
- Veľmi sa to využíva pri stratových kompresiách audio signálov

36

Dane experimenty preukázali, že významne frekvenčné zložky môžu maskovať okolité frekvencie. To sa využíva masívne pri stratovej kompresii zvukových signálov, napr. Mp3, kde dochádza až k 10 násobnej kompresii. To nastane vtedy keď sa neprenesú tie frekvenčné zložky, ktoré sú zamaskované významnými frekvenciami v signále. Proces maskovania je veľmi komplexný a existuje viacero modelov ako určiť maskované frekvencie.

Frekvenčné maskovanie

- Príklad maskovania



Časové maskovanie

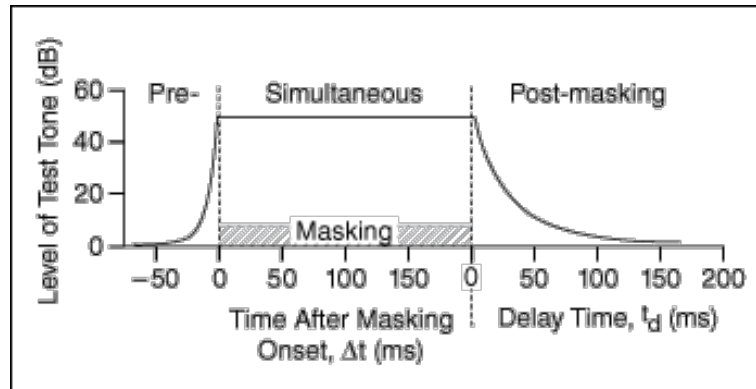
- Intenzívne zvuky môžu zamaskovať menej intenzívne zvuky, ktoré nasledujú alebo dokonca, ktoré nastali skôr
- Rozlišuje sa dopredné a spätné maskovanie
- Dopredné má ale kratšie trvanie
- Táto technika sa ale často nevyužíva v kompresných algoritmoch

38

Okrem dominantného frekvenčného maskovania existuje aj časove maskovanie kedy postupnosť tónov v čase môže ovplyvňovať počuteľnosť buď nasledujúcich ako aj predchádzajúcich tónov. Tento fenomén sa označuje ako časove maskovanie.

Časové maskovanie

- Príklad maskovania



Vnímanie priestoru

- Je možné odhadnúť pozíciu aj vzdialenosť
- Rozdiel vnímanej hlasitosti medzi ušami
 - Nie je presný, len v kombinácii s ostatnými aspektmi
- Časové oneskorenie
 - Najlepšie pulz a vyššie frekvencie
 - Pr.: Vzdialenosť uší 15 cm ale prejde až 23 cm okolo hlavy, teda dorazí o 0,68ms neskôr
 - Ak sú dva zdroje zvuku a posun je menší ako 15ms, smer zvuku je vnímaný, že leží na spojnici oboch zdrojov

40

Zvukové signály je možné využiť aj na lokalizáciu prichádzajúceho zvuku vzhľadom na prítomnosť 2 uší. Na to sa dajú využiť rôzne aspekty šírenia zvuku, rozloženia uší, tvaru a veľkosti hlavy atď.. V ďalšom sú rozobrané niektoré z nich, a to aký je ich vplyv a vnímanie.

Vnímanie priestoru

- Rozdiel fáz signálu medzi ušami
 - Najlepšie stredné tóny
 - Hlboké majú relatívne veľmi malý fázový posun
 - Vysoké, posun môže byť príliš veľký „ďalšia perióda“
- Iná farba zvuku pre ľavé a pravé ucho
 - Rozdielna cesta zvuku k jednému a druhému uchu
 - Prenosová funkcia hlavy