

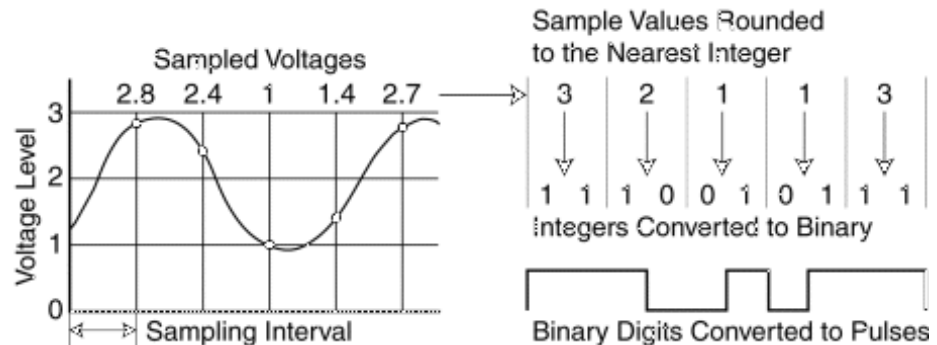
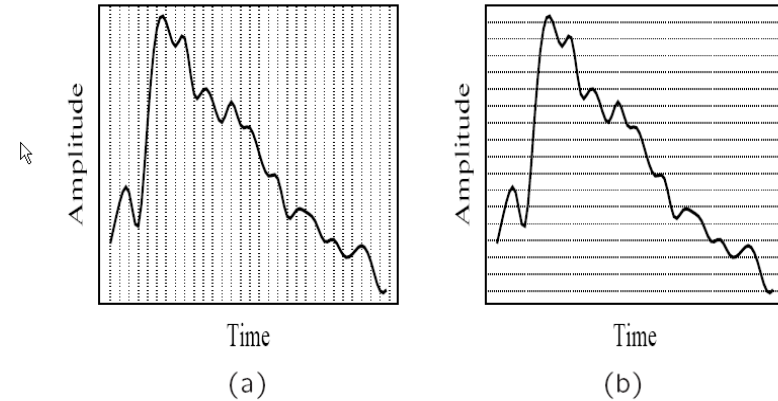
Stratová kompresia údajov

kompresia zvukovej informácie

ÚINF/KMÚ ZS 2019/20
doc. RNDr. Jozef Jirásek, PhD.

Digitalizácia zvuku

- konverzia analógového signálu na postupnosť čísel
- snímkovanie a kvantizácia
- Nyquist – snímkovanie musí byť aspoň dvakrát častejšie ako šírka pásma, ktoré kódujeme
- pre binárne kódovanie
 - telefón 8 bitov
 - CD 16 bitov
- PCM (Pulse Code Modulation)





ITU-T G.711 (telefónne siete)

PCM - 8000 snímok za sekundu, 12-bitová kvantizácia (signál v rozsahu 300 – 3400 Hz)

- uniformná kvantizácia $R = 12 \text{ b} * 8000 / \text{s} = 96 \text{ kb/s}$
- neuniformná 8-bitová kvantizácia log. kompresorom = 64 kb/s

A-law (Európa) $A = 87,6$

$$F(x) = \text{sign}(x) \begin{cases} \frac{A|x|}{1 + \ln(A)} & |x| < 1/A \\ \frac{1 + \ln(A|x|)}{1 + \ln(A)} & \frac{1}{A} \leq |x| \leq 1 \end{cases}$$

μ -law (Amerika , Japonsko) $\mu = 255$

$$F(x) = \text{sign}(x) \frac{\ln(1 + \mu|x|)}{\ln(1 + \mu)} \quad |x| \leq 1$$



ITU-T G.721, G.726

ADPCM 32 kb/s (16, 24, 40 kb/s)

2, 3, 4, 5-bitové rozdielové snímky od A-law (μ -law)

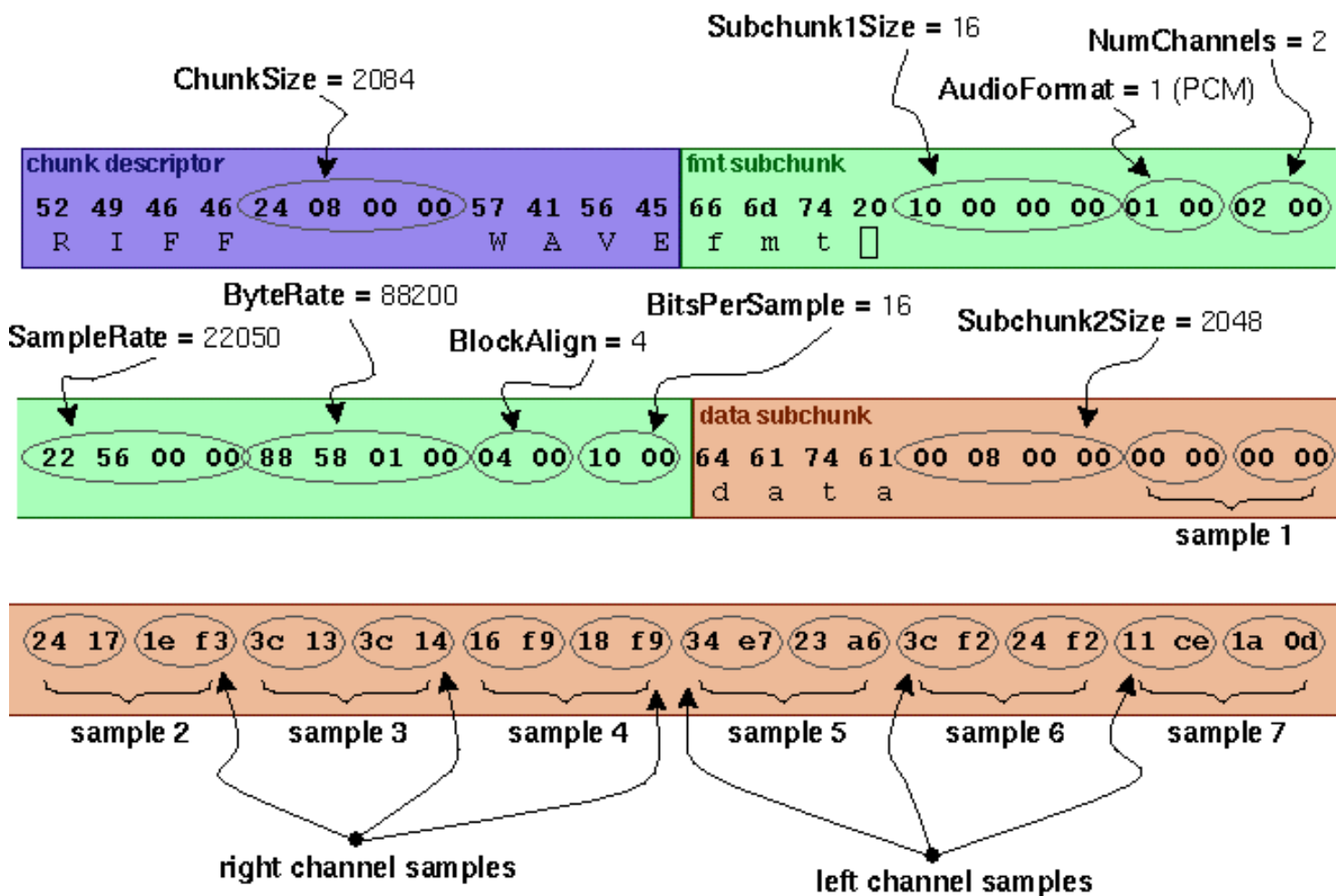
- backward adaptive kvantizátor (podobný Jayantovmu s logaritmickou škálou)
- adaptívny prediktor podľa posledných dvoch rekonštruovaných hodnôt a posledných šiestich kvantizovaných diferencií

$$p_k = \sum_{i=1}^2 a_i^{(k-1)} \hat{x}_{k-i} + \sum_{i=1}^6 b_i^{(k-1)} \hat{d}_{k-i}$$

(rozdielne parametre predikcie pre zvuk a pre reč)

Waveform Audio File Format .WAV

- prekladané (interleaved) viackanálové snímky - bez straty



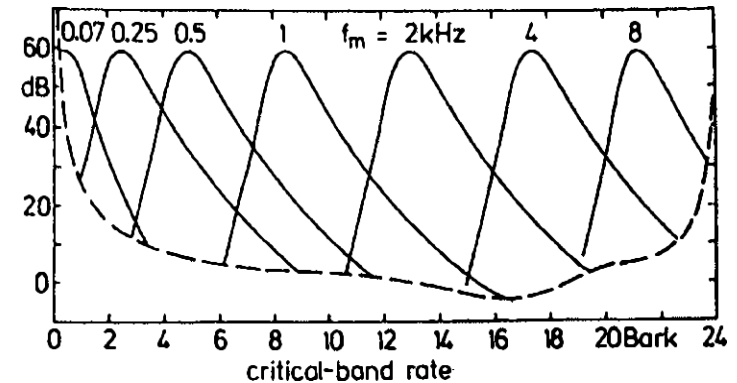
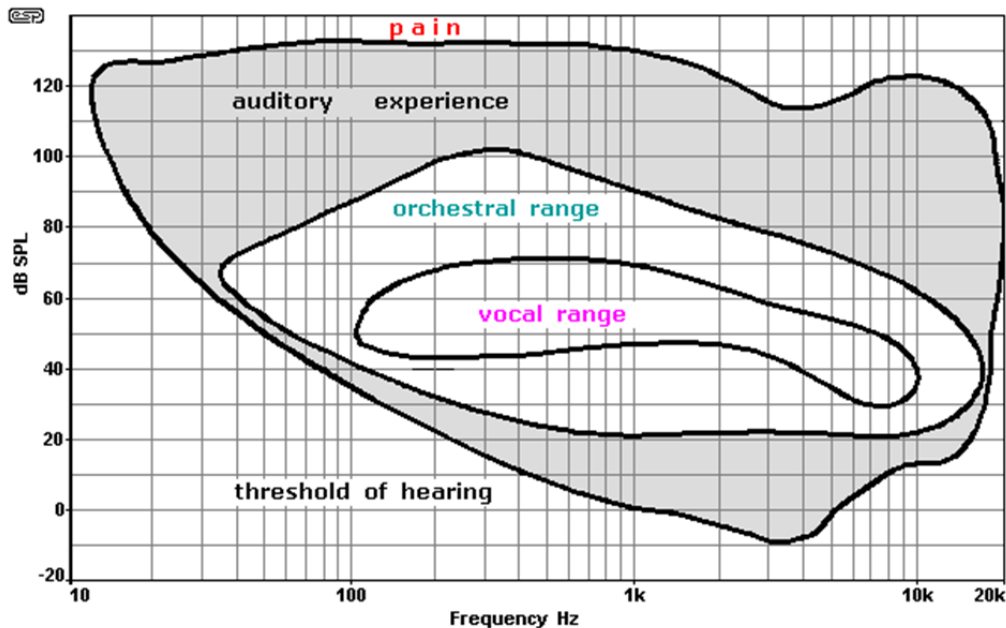


FLAC – Free Lossless Audio Codec

- lineárna predikcia – fixed, FIR linear
- Golomb-Rice kódovanie
- (RLE pre bloky s rovnakou hodnotou – pauzy)
- .flac .mkv .oga
- kompresia 40 – 50 %

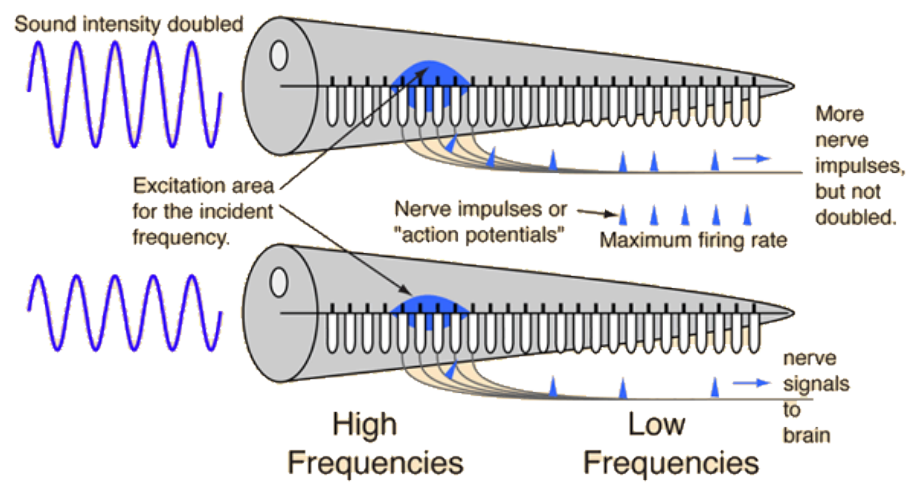
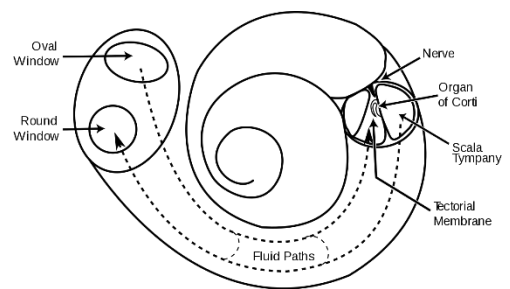
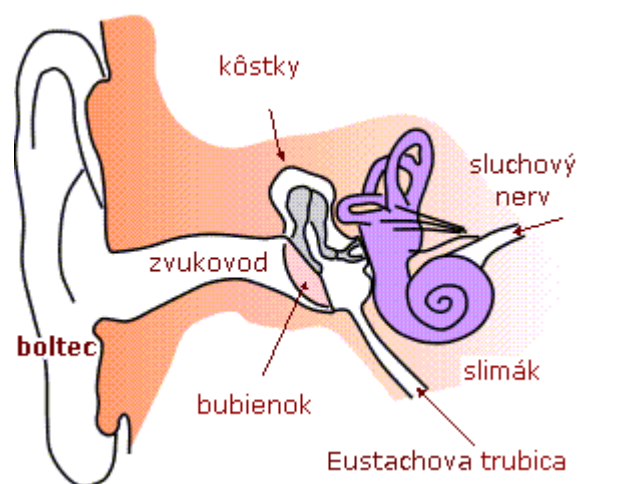
Psychoakustika

- chyby pri vnímaní zvuku (obrazu) človekom
- ohraničenie pásma 20 Hz – 20 kHz
- najlepšie vnímanie 2 – 4 kHz
- fyziologicky 25 pásem (Bark) – širšie pre vyššie frekvencie



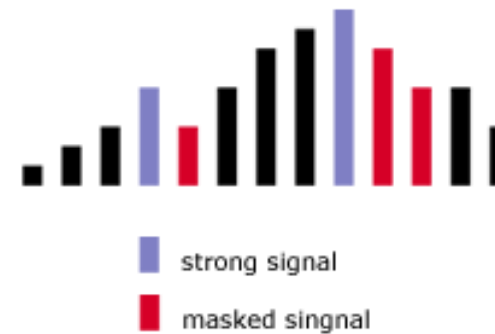
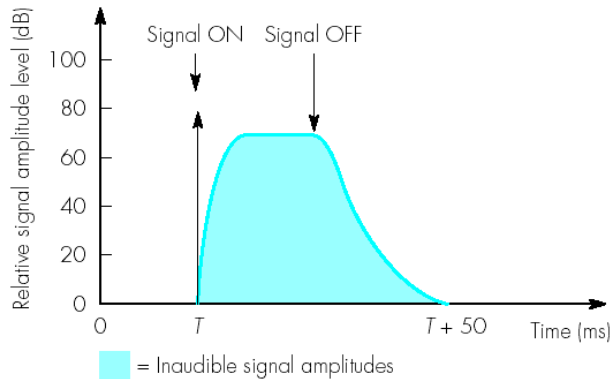
Ucho

- rovnováha + akustické vnemy
- slimák - Cochlea

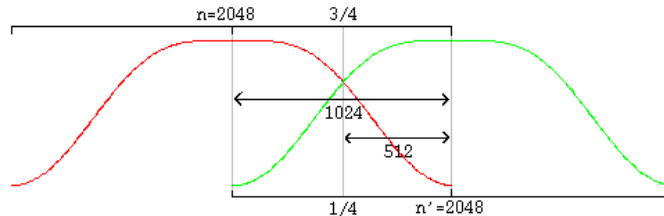
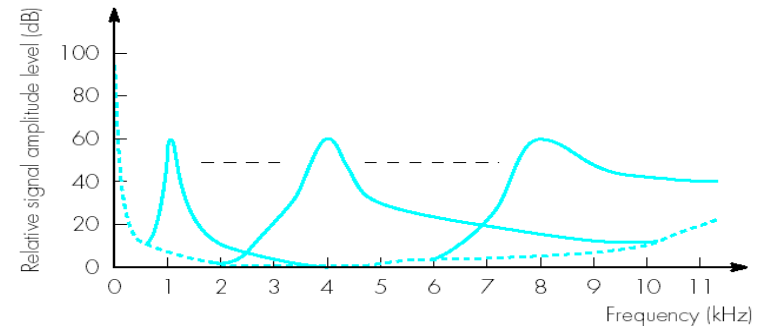


Perceptuálne kódovanie

- maskovanie slabých signálov silnejšími, dozvuk



- MDCT modifikovaná DCT s prekrývaním frekvencií



MPEG (Moving Picture Expert Group) Audio



- MPEG-1: 1,5 Mb/s (1,2 Mb/s video, 0,3 Mb/s audio)
 - kompresný pomer 2.7 – 24, snímkovanie 32, 44.1 a 48 kHz
- 16 b stereo 48 KHz – redukcia na 256 kb/s nie je pozorovateľná
- kanálové režimy
 - Monophonic - single audio channel
 - Dual-monophonic - two independent channels, e.g., English and French
 - Stereo - for stereo channels that share bits, but not using Joint-stereo coding
 - Joint-stereo - takes advantage of the correlations between stereo channels
- rozdelenie na 32 podpásiem PQT (polyphase quadrature filter – podobný ako QMF pre viacero pásem), v každom MDCT
- ak je intenzita v podpásme nižšia ako prah maskovania – nekóduje sa, inak sa vypočíta počet bitov na kódovanie, aby kvantizačný šum bol nižší ako maskovanie (6 dB / bit)

Príklad



Pásmo	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
dB	0	8	12	10	6	2	10	60	35	20	15	2	3	5	3	1

- úroveň 60 dB v 8 pásme maskuje úrovňou 12 dB pásmo 7 a úrovňou 15 dB pásmo 9
- úroveň signálu v pásme 7 je 10 dB (< 12 dB) – nekóduje sa
- v pásme 9 je úroveň 35 dB (> 15 dB) – kódujeme len úrovne nad maskou – namiesto 6 bitového kódu stačí 4-bitový (šetríme 2 bity = 12 dB)

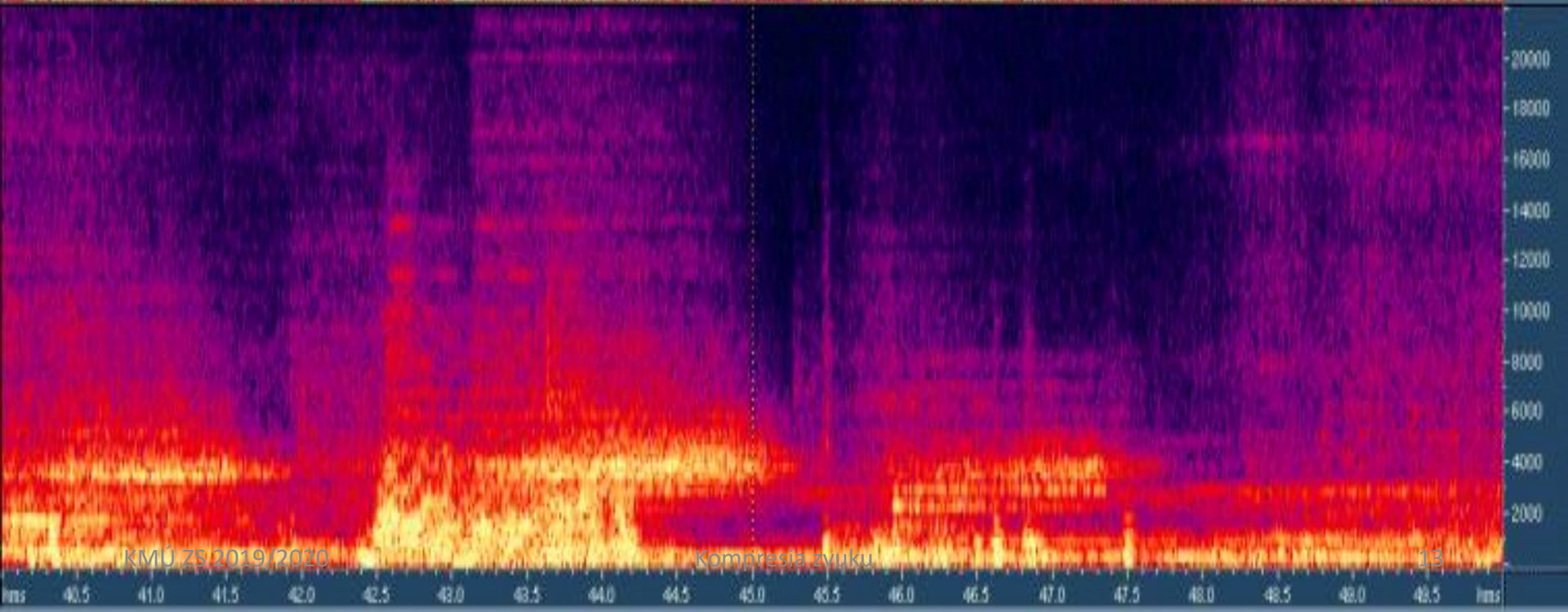
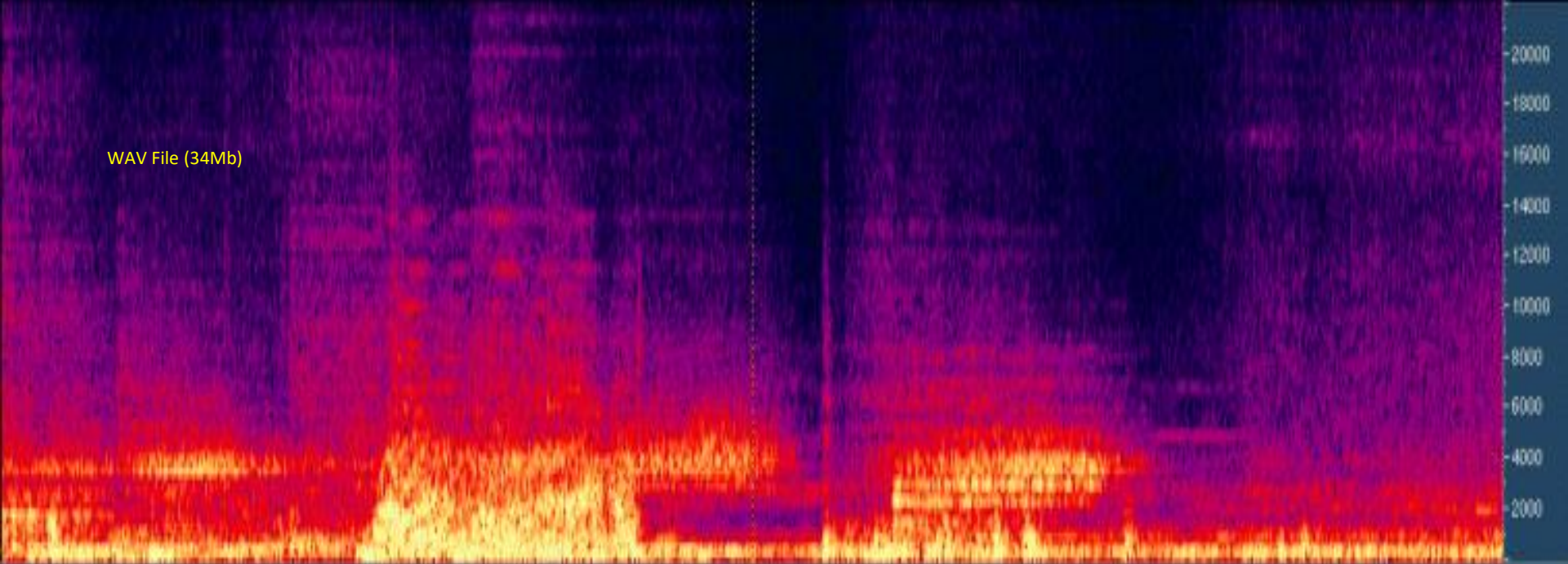


MPEG Audio

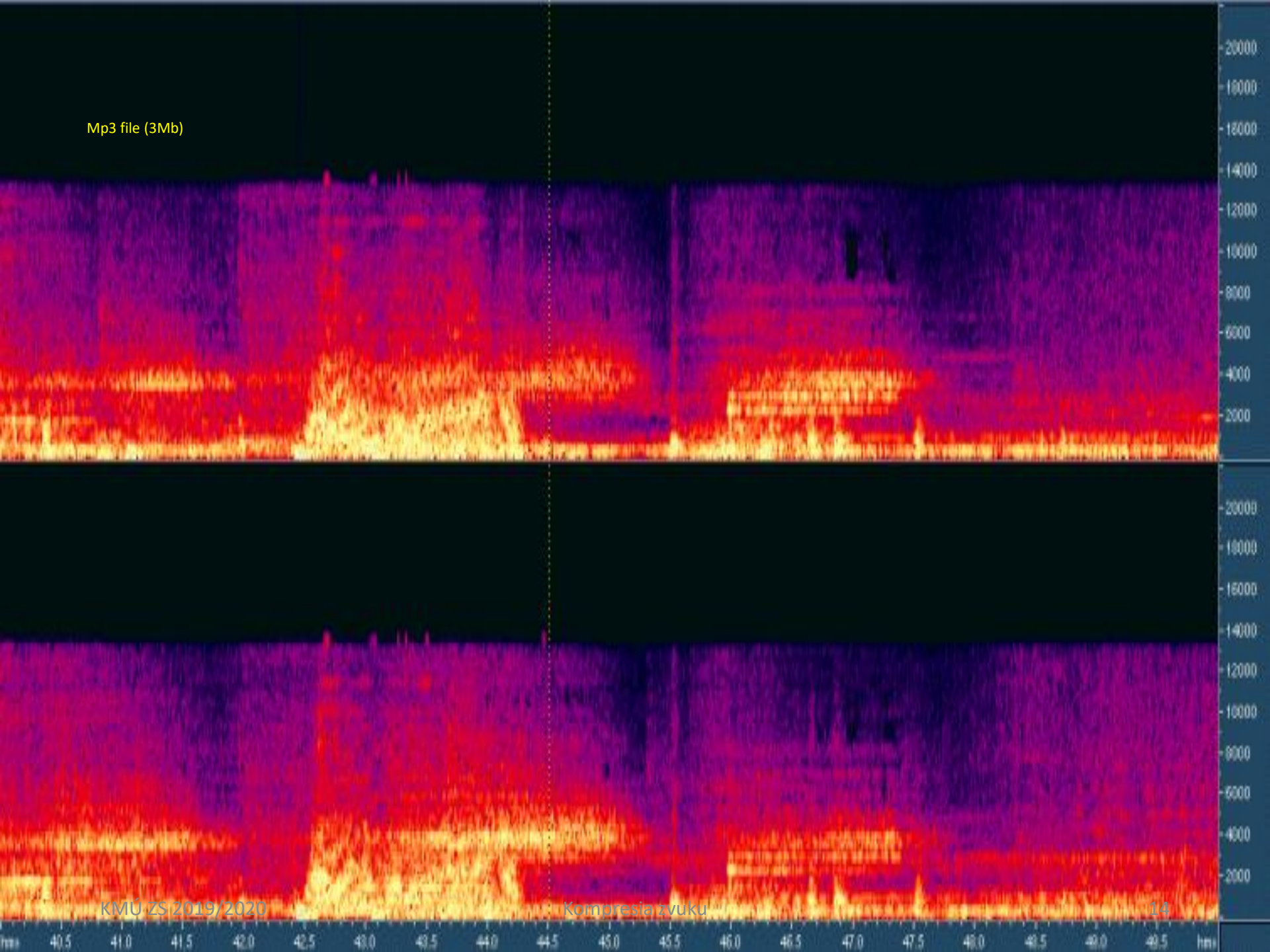
- Layer 1: konštantné šírky podpásem, frekvenčné maskovanie
- Layer 2: 3 filtre (before, current, next) (1152 snímok/rámeček) rôzne škálovanie
- Layer 3 (MP3): nerovnaké frekvenčné pásma, stereo redundancia (vysoké frekvencie rovnako, nízke priemerom a rozdielom L a P kanála), Huffman

Layer	Target Bit-rate	Ratio	Quality at 64 kb/s	Quality at 128 kb/s	Theoretical Min. Delay
Layer 1	192 kb/s	4:1	---	---	19 ms
Layer 2	128 kb/s	6:1	2.1 to 2.6	4+	35 ms
Layer 3	64 kb/s	12:1	3.6 to 3.8	4+	59 ms

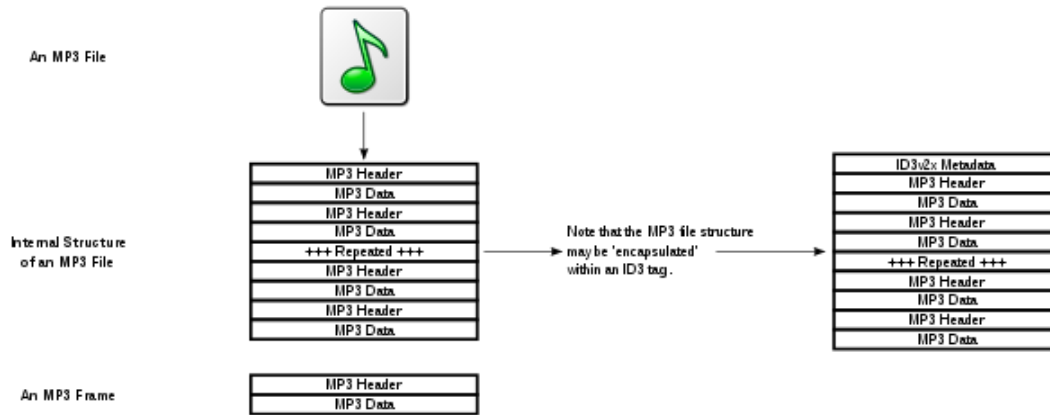
WAV File (34Mb)



Mp3 file (3Mb)



MP3 file format



Example MP3 Header

F F F B A 0 4 0

Colour-coding shows binary bit mapping to hex values below

Detail of an MP3 Header

Bits	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	
Binary	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0
Hex	F			F			F			B			A			0			4			0											
Meaning	MP3 Sync Word												Version	Layer	Error Protection	Bit Rate				Frequency	Pad. Bit	Priv. Bit	Mode	Mode Extension (Used With Joint Stereo)		Copy	Original	Emphasis					
Value	Sync Word												1 = MPEG	01 = Layer 3	1 = No	1010 = 160				00 = 44100 Hz	0 = Frame is not padded	Unknown	01 = Joint Stereo	0 = Intensity Stereo Off	0 = MS Stereo Off	0 = Not Copyrighted	0 = Copy Of Original Media	00 = None					



Vorbis

- projekt pre audio kodeky
- MDCT bloky 64 – 8192 snímok
- lineárne filtre (po častiach)
- vektorová kvantizácia a Huffmanovo kódovanie
- viacprechodové kódovanie, korelácia medzi kanálmi



Dolby Digital

- 32, 44.1, 48 kHz 5+1 kanálov 32 – 640 kb/s
- 1536 snímkové rámce, MDCT po 512 blokoch
- kombinácia kanálov s podobným obsahom
- súčtový a diferenčný kanál pre stereo
- dynamická alokácia bitov pre kódovanie



AAC (Advanced Audio Coding)

- vylepšenie MP3 v MPEG-2/4 špecifikácii
- frekvencia snímkovania 8 – 96 kHz
- MDCT po blokoch 2048 resp. 256
- až 48 kanálov
- kompanderová kvantizácia
- Huffman/arithmetic + VQ

- SBR – Spectral Band Replication
- optimalizácia ukladania vysokých frekvencií (obálky, rekonštrukcia z nízkych frekvencií) 25 – 50 % redukcia DR
- HE-AAC resp. mp3PRO



audio médiá

- SACD (super audio CD), DVD-A (DVD audio)
- DSD (Direct Stream Digital)

	16-, 20- or 24-bit (DVD-A)					
	44.1 kHz	48 kHz	88.2 kHz	96 kHz	176.4 kHz	192 kHz
Mono (1.0)	Yes	Yes	Yes	Yes	Yes	Yes
Stereo (2.0)	Yes	Yes	Yes	Yes	Yes	Yes
nStereo (2.1)	Yes	Yes	Yes	Yes	No	No
Stereo + mono surround (3.0 or 3.1)	Yes	Yes	Yes	Yes	No	No
Quad (4.0 or 4.1)	Yes	Yes	Yes	Yes	No	No
3-stereo (3.0 or 3.1)	Yes	Yes	Yes	Yes	No	No
3-stereo + mono surround (4.0 or 4.1)	Yes	Yes	Yes	Yes	No	No
Full surround (5.0 or 5.1)	Yes	Yes	Yes	Yes	No	No

Data Rate vs. Distortion



Quality	Sample Rate (KHz)	Bits per Sample	Mono/Stereo	Data Rate (uncompressed) (kB/sec)	Frequency Band (KHz)
Telephone	8	8	Mono	8	0.200-3.4
AM Radio	11.025	8	Mono	11.0	0.1-5.5
FM Radio	22.05	16	Stereo	88.2	0.02-11
CD	44.1	16	Stereo	176.4	0.005-20
DAT	48	16	Stereo	192.0	0.005-20
DVD Audio	192 (max)	24 (max)	6 channels	1,200.0 (max)	0-96 (max)

- CD 1,411 Mb/s
- MP3 96, 128, 160 kb/s po kompresii
- VoIP 5,3 – 13 kb/s po kompresii

Ďakujem za pozornosť.

jozef.jirasek@upjs.sk