

# Moderní multimedialní elektronika (U3V)

Prezentace č. 13

## Moderní kompresní formáty pro přenosné digitální audio



Ing. Tomáš Kratochvíl, Ph.D.

Ústav radioelektroniky, FEKT VUT v Brně

# Program prezentace

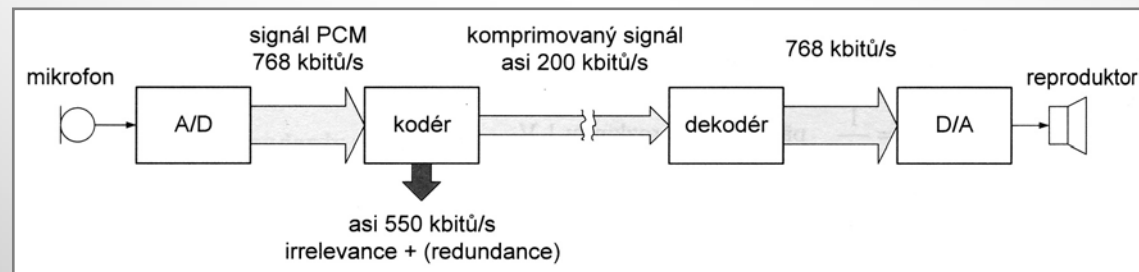
- ❑ Princip redukce datového toku v oblasti audia
- ❑ Perceptivní kódování a maskování zvuku
- ❑ MPEG standard pro kódování audia
- ❑ Kompresní formáty MP3, Atrac, AAC, OGG, WMA, VQF
- ❑ Praktické ukázky ovládání přenosných audio přístrojů
- ❑ Praktické ukázky kvality reprodukce komprimovaného zvuku v laboratoři Nízkofrekvenční elektroniky UREL FEKT

# Princip redukce datového toku v oblasti audia

Výrazné zmenšení bitové rychlosti se dosáhne kompresí dat (zvuku), tedy potlačením dat, která jsou redundantní (nadbytečná) a irrelevantní (zbytečná).

Redundance – příkladem je stereofonní signál se shodnou informací v obou kanálech, která je nadbytečná.

Irrelevance – je založená na vlastnostech lidského vnímání zvuku (omezení slyšitelnosti na určitou frekvenční oblast, maskování slabších signálů).

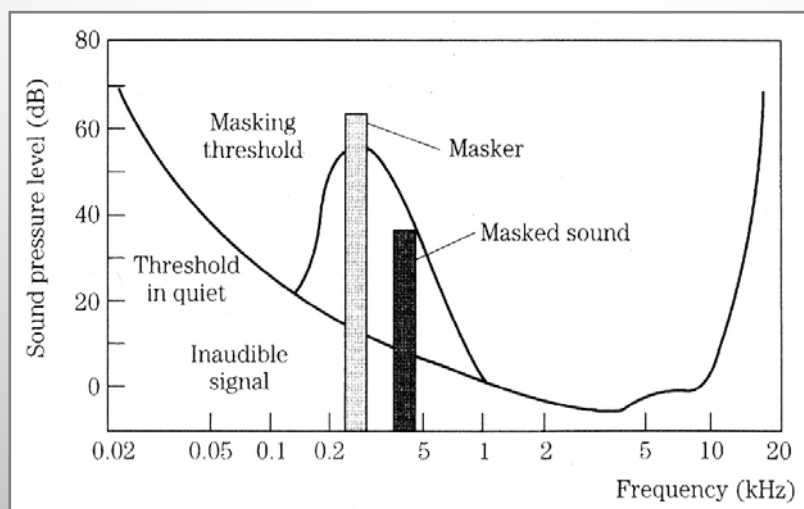


Snížení bitové rychlosti při digitálním přenosu zvuku

Zdrojové kódování PCM, kvantování na 16b a se vzorkovací frekvencí 48 kHz. Kvantizační šum je maskován vyšší úrovní užitečného signálu. Je zbytečné (irrelevantní) přenášet i silné akustické signály stále stejným a velkým počtem kvantizačních stupňů a v určitých frekvenčních pásmech se počet zmenšuje.

## Perceptivní kódování - princip maskování slyšitelných tónů

Při poslechu není vnímán jeden sinusový tón, ale směs různých frekvencí – tónové (sinusové průběhy) a netónové (šumy, údery apod.). Při současném trvání zvuků může silnější zvuk potlačovat slabší, i když nemají stejnou frekvenci – silnější jej při určitém akustickém tlaku maskuje.



Maskovací zvuk nebo šum zvyšuje práh slyšitelnosti v lokální oblasti (kritickém kmitočtovém pásmu) a maskuje sousední slabší zvuk.

## Temporální maskování - premasking a postmasking

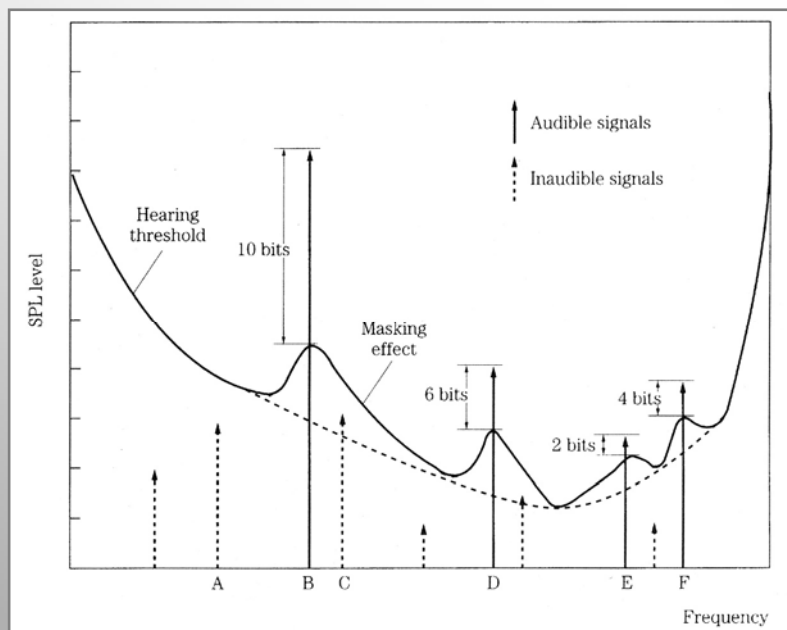
Maskování funguje také i v případě, že maskovaný krátkodobý signál přichází až po ukončení maskujícího zvuku (v době do 10 ms, jinak maskování dále slábne a po 200 ms zcela zaniká). Rovněž může být maskován krátký zvukový impuls, následuje-li po něm maskující signál (v době do 5 ms) trvajícím déle.

Základem temporálního maskování je předpoklad, že mozek vnímá zvuk jako integraci v čase a zpracovává ji sluchovým orgánem v časových intervalech.

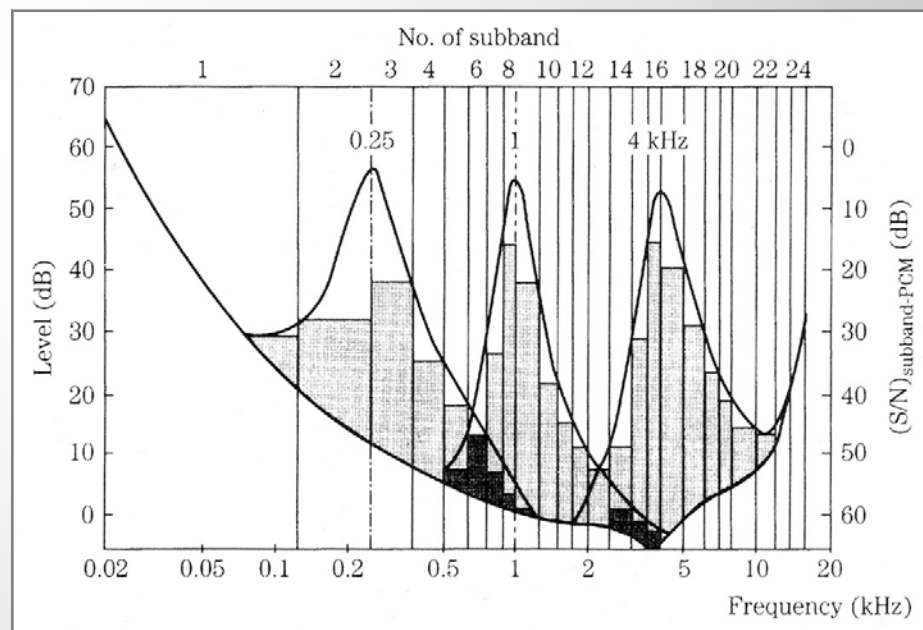
## Maskování šumu zvukovým signálem

Zvuk je tvořen harmonickými signály a křivky maskovacích prahových hodnot vymezují rozhraní, pod kterým nejsou ostatní jiné zvuky uvažovaných frekvencí a odečtených akustických hladin slyšitelné.

Tato vlastnost se využívá pro maskování šumu tak, že šumové spektrum je tvarováno podle nejsilnějších složek zvukového signálu. Z proměnné hladiny šumu vychází požadavek na minimální počet bitů pro kvantování.



Přidělení počtu bitů pro kvantování na jednotlivých frekvencích a v jednotlivých pásmech



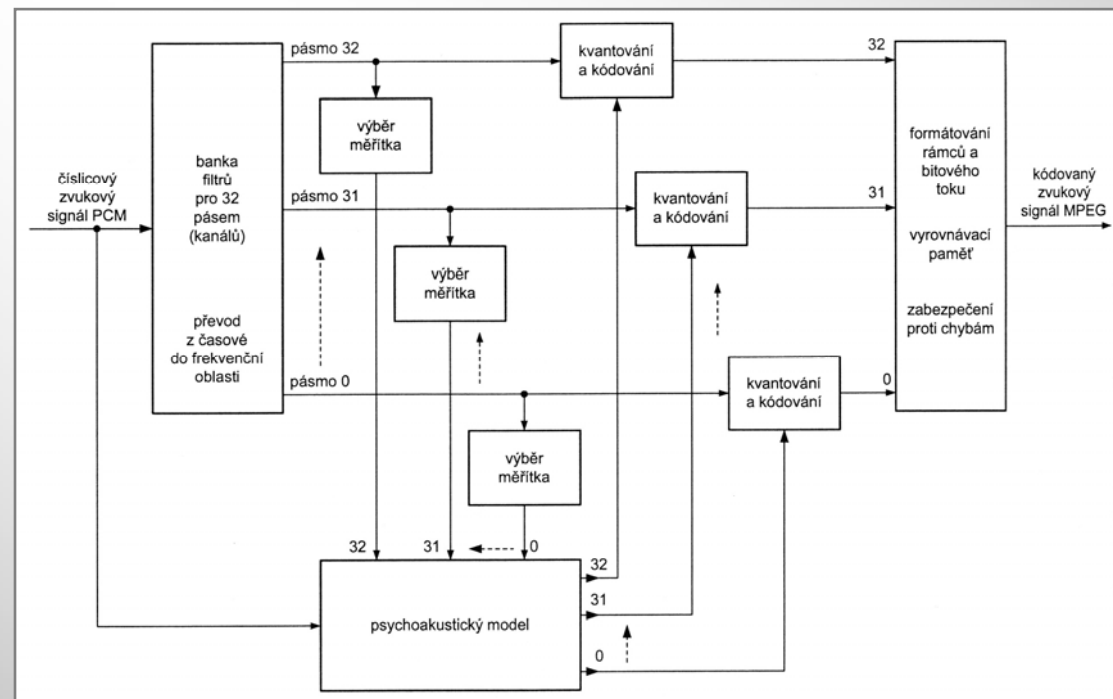
Maskování kvantizačního šumu v dílcích frekvenčních pásmech 3 tónovými frekvencemi

# MPEG standard pro kódování audia

V bance filtrů se číslicový signál převede z časové oblasti do frekvenční oblasti. Vytváří se základní skupiny sériových dat, které se nazývají datové rámce. Soubor frekvenčních složek každého rámce se rozdělí do 32 stejně širokých dílčích frekvenčních pásem. V každém pásmu se vybere vzorek s maximální amplitudou a podle něj se určí činitel měřítka, který má rozhodující úlohu pro stanovení maskovacích prahů v jednotlivých pásmech.

Měřítka jsou srovnány s maskovacími prahy psychoakustického modelu (který respektuje vlastnosti lidského sluchu) a stanoví se hladina maskovaného kvantizačního šumu a přidělí se počet bitů na kvantování každého pásma.

MPEG audio kodér pro redukci bitového toku

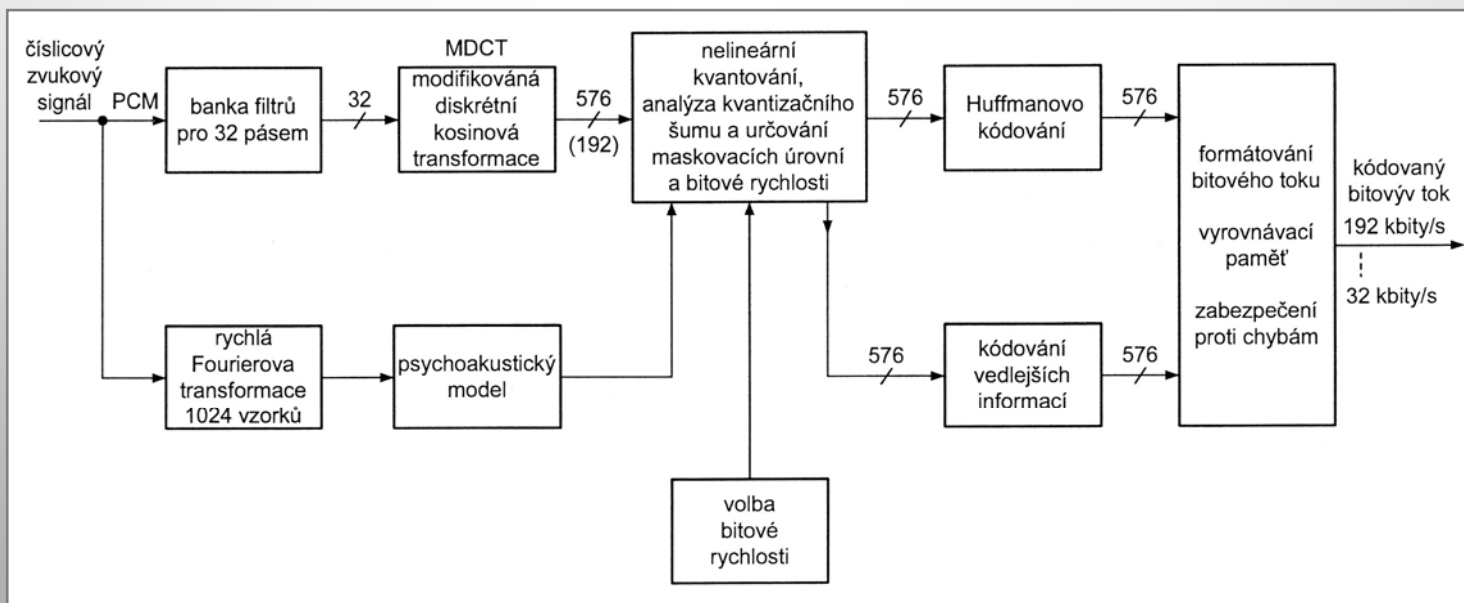


## Kodér MPEG-1 Layer III (MP3)

Hybridní banka filtrů rozdělí signál do 32 frekvenčních pásem. Jemného spektrálního rozlišení se dosáhne použitím MDCT (modifikované diskrétní kosinovy transformace) zapojené v kaskádě s bankou filtrů.

Příliš úzká frekvenční filtrace však může způsobit chyby výstupního signálu, proto se celkový počet spektrálních koeficientů podle povahy vstupního signálu může zmenšit až na 6 v každém pásmu. Časové rozlišení je 24 ms při 48 kHz.

Kvantovací obvod je nelineární a navíc je zde Huffmanovo kódování, které snižuje objem datového toku.



MPEG-1 Layer III - kodér

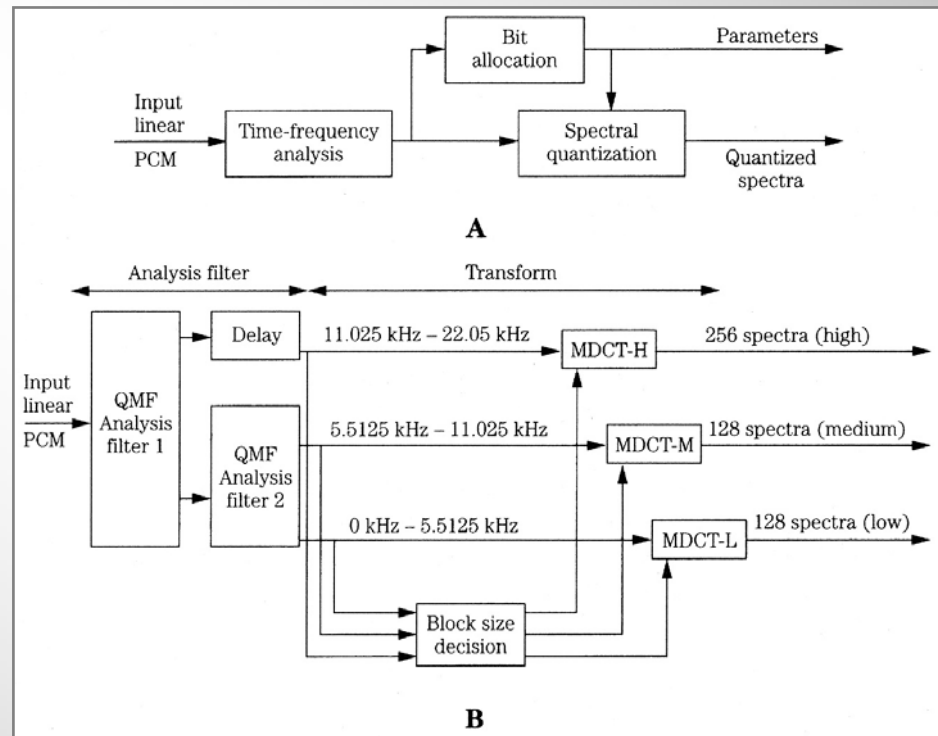
# ATRAC (Sony Minidisc) datová redukce

## Kodér ATRAC (Adaptive TRansform Acoustic Coding)

Perceptivní kódování s bitovým tokem 292 kbit/s, standardní vzorkovací kmitočet 44,1 kHz komprese cca 1/5.

Kodér rozděljuje vstupní audio vzorky do časových bloků, které rozděljuje do 3 frekvenčních subpásem, ve kterých provádí MDCT. Celkově vznikne 512 frekvenčních vzorků v kritických pásmech a provádí se maskování kvantizačního šumu a přidělení počtu bitů pro kvantizaci.

ATRAC se liší od předchozích aplikací kombinací subpásmových a transformačních technik s psychoakustickým modelem. Amplitudové a časové rozlišení se mění spolu s dynamikou signálu.



ATRAC a) kodér, b) detail časově-frekvenční analýzy.

# Kompresní formáty v oblasti audio na PC

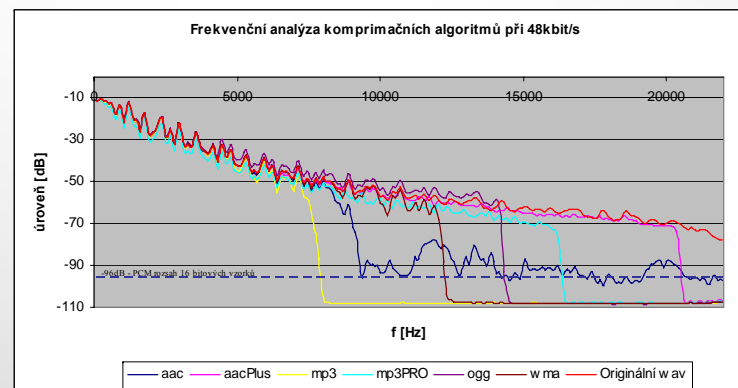
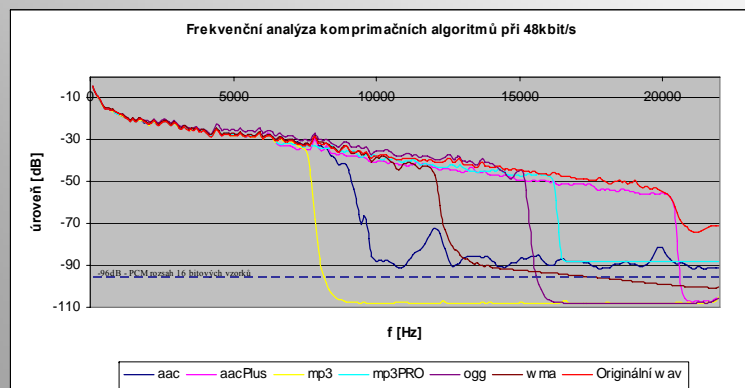
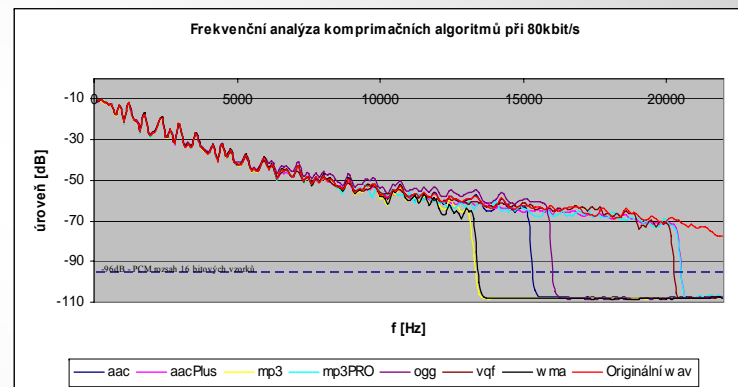
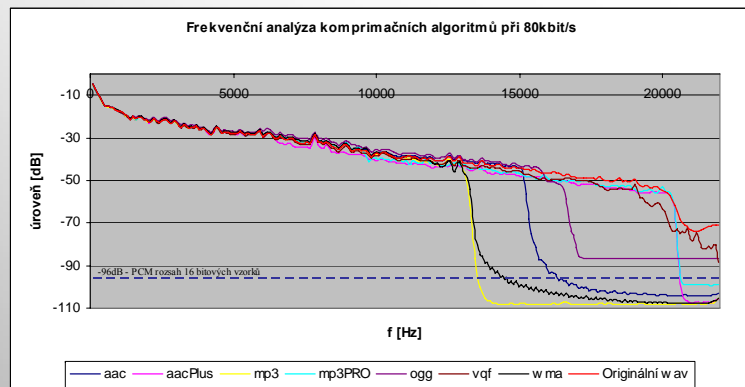
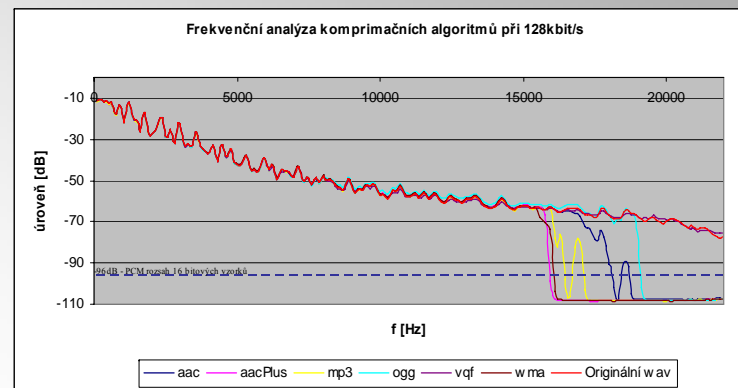
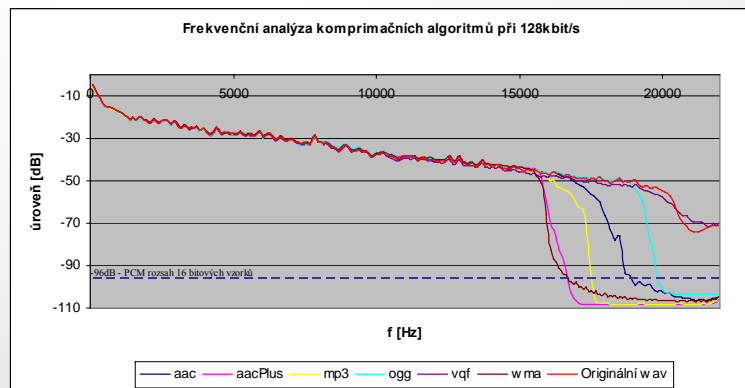
- ❑ MP3 – standardizován v MPEG-1 Layer 3, nejrozšířenější formát na PC, podporuje CBR (Constant Bit Rate) a VBR (Variable Bit Rate), kvalita CD odpovídá 128 kbit/s.
- ❑ MP3 Pro – využívá SBR (Spectral Band Replication), doplňkové kódování vyšších spektrálních pásem, kompatibilní s MP3 při přehrávání.
- ❑ AAC – standardizován v MPEG-2 Advance Audio Coding, efektivní metoda pro kódování vícekanálového zvuku, typicky pro 5.1 zvukový doprovod
- ❑ AAC Plus – rozšiřuje formát AAC o SBR v oblasti vysokých kmitočtů (10-20 KHz)
- ❑ OGG – kodek až pro 255 kanálů, použití VBR (Variable Bit Rate), nepodporuje rámce a zabezpečení, jednoduchá konverze při snížení bitové rychlosti.
- ❑ WMA – Microsoft Windows Media Audio, splňuje podmínky pro licencování skladeb, jednotlivé verze podporují bezztrátové kódování (WMA Pro), kódování vícekanálového zvuku, hlasu (WMA Voice), podporuje CBR a VBR.
- ❑ VQF – Vector Quantization File, podobné vlastnosti jako MP3, propracovanější kódování a vyšší kvalita při shodném bitovém toku.



# Srovnání jednotlivých komprimačních algoritmů

Elektronická hudba – Genesis, Brazilian

Klasická hudba – Mozart, Figarova svatba



## Praktické ukázky reprodukce a ovládání



Flash přehrávač  
iRiver iFP-1095



HDD Audio/Video přehrávač  
iRiver PMP-140



CD/MP3/DivX přehrávač  
iRiver iMP-1100