



## Semestrální práce

### Kompresní algoritmy v bezztrátových zvukových formátech.

*Studijní program:*

*Studijní obor:*

*Autor práce:*

*Vedoucí práce:*

N0613A140028 – Informační technologie

N0613A140028IS – Inteligentní systémy

**Bc. David Šafařík**

doc. Ing. Otto Severýn, Ph.D.

Liberec 2022

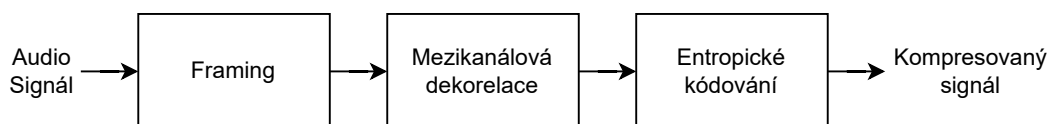
# Obsah

<b>1</b>	<b>Úvod</b>	<b>1</b>
<b>2</b>	<b>Kompresní algoritmy ve vybraných bezztrátových zvukových formátech</b>	<b>3</b>
2.1	Free Lossless Audio Codec . . . . .	3
2.2	MPEG-4 Audio Lossless Coding . . . . .	4
2.3	Apple Lossless . . . . .	5
2.4	Monkey's audio . . . . .	5
2.5	SHORTEN . . . . .	6

## 1 Úvod

Bezeztrátová komprese je výhodná z hlediska toho, že po dekomprimaci nedochází ke ztrátě informace, data jsou tak stejná jako před komprimací. Nicméně účinnost není tak vysoká jako u ztrátové komprese. U ztrátové komprese dochází ke ztrátě určitého množství dat. V tomto případě je využíváno nedokonalosti lidských smyslů, jako třeba že lidské ucho neslyší celé spektrum zvukového signálu stejně. [1]

Existuje mnoho způsobů bezeztrátové komprese audio signálu, většina kompresních algoritmů využívá lineární prediktivní kódování (LPC). Kodér lze považovat za základní operaci většiny bezztrátových kompresních algoritmů, viz obr. 1.1, kde blok mezikanálové dekorelace obsahuje predikční model spolu s kvantizační operací pro každý z kanálů audio signálu. V této části se vždy vyskytne chyba v odhadu, vzniká tak ztráta informace, a proto následuje blok entropického kódování. Tento blok vynahrazuje ztráty z předchozího bloku, a zajišťuje tak bezeztrátovou kompresi. To je prováděno zakódováním rezidua chyby predikčního modelu. Blok entropického kódování tak sečte zakódované reziduum chyby s výsledkem predikčního modelu a tím tedy zajišťuje bezeztrátovou kompresi. [2]



Obrázek 1.1: Obecné schéma bezztrátového zvukového kodéru [3]

Predikční model zde působí jako ztrátový dekodér, který rekonstruuje původní signál s co nejmenší odchylkou. Čím menší je odchylka, tím je účinnější entropické kódování, a tím je menší kompresovaný signál. Obrázek 1.2 popisuje běžnou strukturu predikčního modelu, ve které je výstup z predikčního modelu reziduum chyby odhadu, které je zakódováno spolu s koeficienty predikčního modelu. [2]

Princip lineárního prediktivního kódování spočívá v tom, že vzorek zvukového signálu  $x(n)$  lze predikovat jako lineární kombinaci předchozích vzorků  $x(n - k)$  [4] podle vzorce:

$$\hat{x}(n) = \sum_{k=1}^K a_k x(n - k) \quad (1.1)$$

Kompresovaný signál je pak typicky spočítán jako rozdíl predikovaného signálu a původního signálu podle vzorce:

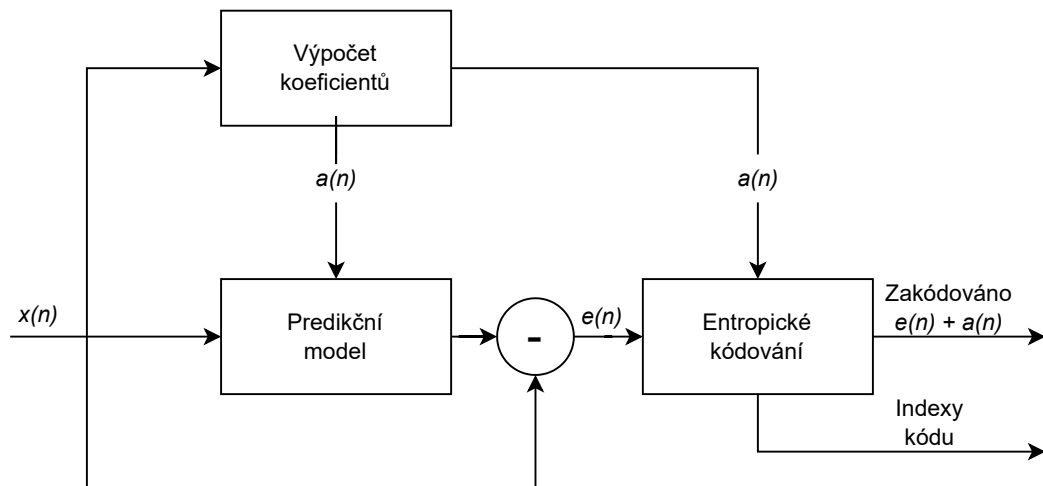
$$e(n) = x(n) - \hat{x}(n) \quad (1.2)$$

Výčet některých bezztrátových zvukových formátů:

- Adaptive Transform Acoustic Coding (ATRAC),
- Apple Lossless (ALAC – Apple Lossless Audio Codec),
- Audio Lossless Coding (also known as MPEG-4 ALS),
- Direct Stream Transfer (DST),
- Free Lossless Audio Codec (FLAC),
- Meridian Lossless Packing (MLP),
- Monkey's Audio (Monkey's Audio APE),
- MPEG-4 SLS (also known as HD-AAC),
- Original Sound Quality (OSQ),
- RealPlayer (RealAudio Lossless),
- Shorten (SHN),
- TTA (True Audio Lossless),

- WavPack (WavPack lossless),
- WMA Lossless (Windows Media Lossless).

Dále v práci budou přiblíženy algoritmy pouze některých z nich.



Obrázek 1.2: Obecné struktura predikčního modelu s entropickým kódováním [2]

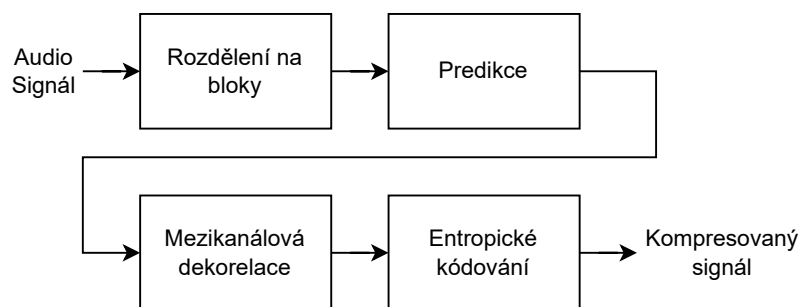
## 2 Kompresní algoritmy ve vybraných bezztrátových zvukových formátech

V této kapitole jsou popsány algoritmy komprese vybraných bezztrátových zvukových formátů.

### 2.1 Free Lossless Audio Codec

Kodek známý pod zkratkou FLAC<sup>1</sup> je otevřený kodek, který pochází z roku 2001. Průběh komprese je naznačen na obrázku 2.1. Kompresní poměr tohoto kodeku je horší než některých jiných bezztrátových kodeků, ale jeho rychlost dekomprese vyvažuje tento nedostatek.

<sup>1</sup><https://xiph.org/flac/>



Obrázek 2.1: Struktura FLAC kodéru [2]

Fáze rozdělení na bloky rozdělí signál na bloky nebo části dané velikosti. Volba velikosti bloků přímo ovlivňuje kompresní poměr. Mezikanálová dekorelace je uskutečněna odstraněním redundance v levém a pravém kanálu. Levý a pravý kanál jsou zakódovány do středového kanálu (průměr kanálů) a do vedlejšího kanálu (rozdíl levého a pravého kanálu). Tím je snížen počet bitů potřebný k uložení signálu. V případě, že by kanály byly dost odlišné je možné tuto fázi přeskočit. Fáze predikce je nutná pro dobrou kompresi a může se lišit část od části vzorků signálu. [2]

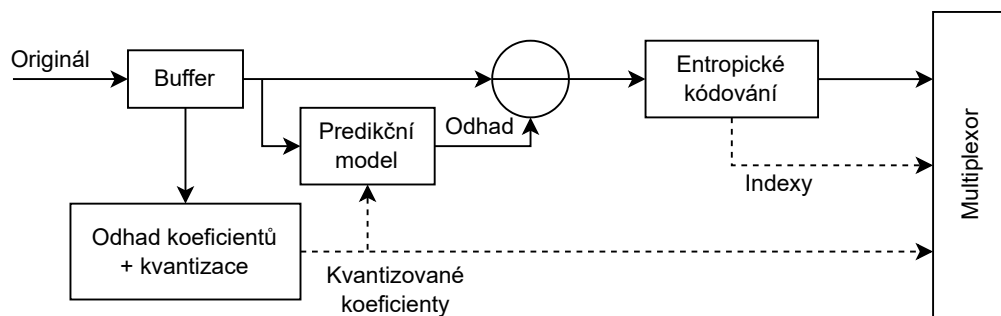
Tento kodek využívá lineární prediktivní kódování pro konverzi zvukových vzorků do tzv. reziduálů, které jsou uloženy pomocí Golom-Ricova kódování. Dále využívá run-length encoding RLE, které je ale efektivní pouze do 8-bitového rozlišení. [1]

## 2.2 MPEG-4 Audio Lossless Coding

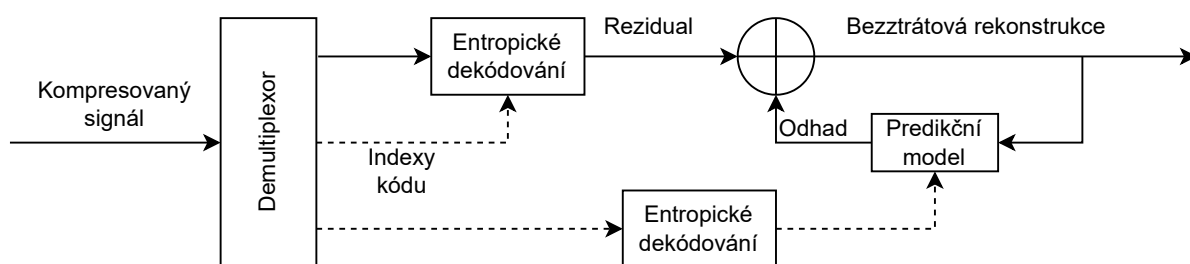
MPEG-4 ALC je jeden z několika MPEG-4 standardů. Kódér tohoto formátu se skládá z pěti bloků (obr. 2.2) (popis těchto bloků je odvozen z úvodní kapitoly):

- Buffer: frame signálu;
- Odhad koeficientů a kvantizace, blok predikčního modelu: mezikanálová dekorelace;
- Entropické kódování: kóduje reziduum chyby pomocí Golomb-Ricova kódování;
- Multiplexor: na výstupu má kompresovaný signál s příslušnou hlavičkou.

Pro dekompresi obsahuje odpovídající dekodér demultiplexor, který rozloží dekódovaný signál a znovu zakóduje reziduum, indexy kódu a kvantizované koeficienty. Viz obr. 2.3. [5]



Obrázek 2.2: MPEG-4 ALC kodér [2]



Obrázek 2.3: MPEG-4 ALC dekodér [5]

## 2.3 Apple Lossless

Apple Lossless kodek (též ALAC nebo ALE) využívá lineární predikci a Golomb-Ricova kódování.

## 2.4 Monkey's audio

Bezeztrátový kodek s rychlým kódováním, ale náročným dekódováním. Podporuje pouze 2 kanály. Účinnost komprese je okolo 50%, což je ale na úkor rychlosti.

Kodek využívá detekci chyb a připojuje cyklický redundantní součet. Ten slouží k bezchybné dekompresi dat.

Při kompresi je využíváno optimálního ukládání kanálů - průměr kanálů a rozdíl kanálů. Poté předpovídá průběh dat pomocí lineární algebry. Předpokládaná hodnota se porovná s původním signálem, na základě čehož se vyhodnotí, zda byla předpověď užitečná či nikoliv. Podle toho se následně předpověď přizpůsobuje nahoru nebo dolů. Pokračuje se výpočtem rozdílu z původního signálu a předpovědi. Snaha je ukládat co nejmenší hodnotu, která se zakóduje pomocí Riceova kódu. [1]

## 2.5 SHORTEN

SHORTEN využívá lineární prediktivní kódování a Huffmanovo kódování pro kompresi zvukového signálu. Zakódovaný signál je pak rozdíl původního signálu a hodnot predikovaných predikčním modelem. [6]

## Bibliografie

1. LENOCH, Václav. *Kompresa zvukových signálů*. Pardubice, 2009. Dostupné také z: <http://hdl.handle.net/10195/34227>. bakalářská práce. Univerzita Pardubice, Dopravní fakulta Jana Pernera.
2. MUIN, Fathiah Abdul; GUNAWAN, Teddy Surya; KARTIWI, Mira; ELSHEIKH, Elsheikh MA. A review of lossless audio compression standards and algorithms. In: *AIP Conference Proceedings*. AIP Publishing LLC, 2017, sv. 1883, s. 020006. Č. 1.
3. HANS, M.; SCHAFER, R.W. Lossless compression of digital audio. *IEEE Signal Processing Magazine*. 2001, roč. 18, č. 4, s. 21–32. Dostupné z DOI: [10.1109/79.939834](https://doi.org/10.1109/79.939834).
4. O'SHAUGHNESSY, D. Linear predictive coding. *IEEE Potentials*. 1988, roč. 7, č. 1, s. 29–32. Dostupné z DOI: [10.1109/45.1890](https://doi.org/10.1109/45.1890).
5. LIEBCHEN, Tilman; MORIYA, Takehiro; HARADA, Noboru; KAMAMOTO, Yutaka; REZNIK, Yuriy A. The MPEG-4 Audio Lossless Coding (ALS) standard-technology and applications. In: *Proc. 119th AES Conv.* 2005, s. 1–14.
6. ROBINSON, Tony. *SHORTEN: Simple lossless and near-lossless waveform compression*. University of Cambridge, Department of Engineering, 1994.