

# EXPERIMENTÁLNÍ ZVUKOVÝ KODÉR

*F. Rund, J. Nováček*

Katedra radioelektroniky K13137, FEL ČVUT Praha

## Abstrakt

Všechny dnes široce rozšířené systémy pro kompresi zvuku vycházejí ze stejných psychoakustických principů a využívají velmi podobnou vnitřní blokovou strukturu. To nám umožňuje vytvoření jediného modulárního systému, který bude mít definován pouze formát dat vstupujících, respektive vystupujících z jednotlivých systémových bloků. Vnitřní struktura daného bloku není přímo definována a závisí pouze na realizovaném kompresním algoritmu.

## 1 Úvod

V dnešní době je komprese zvuku velmi aktuálním problémem. Přes neustále se zvětšující kapacitu záznamových médií je kladen veliký důraz na snižování objemu dat potřebných pro zakódování dané informace. Tento trend postihl i oblast záznamu a přenosu širokopásmových zvukových signálů. Od vzniku dnes bezpochyby nejrozšířenějšího standardu pro kompresi zvuku MP3 (MPEG 1 Layer III) začátkem 90. let 20. století se objevilo mnoho dalších, které více či méně na tento formát navazují.

Tento článek popisuje univerzální modulární systém, pomocí kterého je možné porovnávat a kvalitativně hodnotit jednotlivé kompresní algoritmy z hlediska vnímání zvuku lidským sluchem. Navržený systém umožňuje nejen implementaci, studium a hodnocení stávajících systémů (například kodeky z rodiny MPEG, kodek Ogg Vorbis nebo kodeky ATRAC), ale také implementaci nových algoritmů pro kompresi zvukových signálů založených na nových poznatcích z oblasti lidského slyšení.

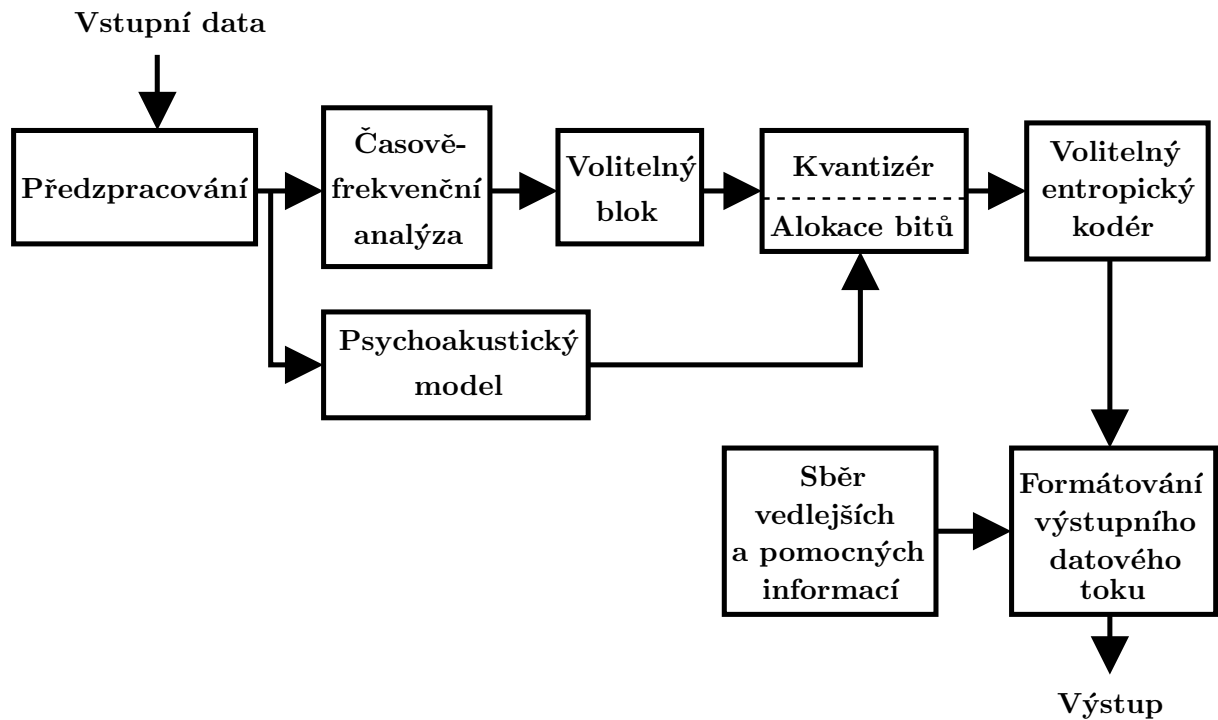
## 2 Univerzální modulární systém

Výchozí myšlenkou celého návrhu je fakt, že velká většina systémů pro kompresi zvukových signálů vychází ze stejných psychoakustických principů a proto využívají velmi podobných vnitřních struktur (viz např. [1], [2], aj.) Na základě analýzy těchto systémů jsme navrhli blokové schéma podle obr. 1. Toto blokové schéma bylo navrženo tak, aby bylo maximálně univerzální a umožňovalo implementaci co nejvyššího počtu kompresních algoritmů. Jedním z hlavních cílů navrženého univerzálního systému je možnost porovnání algoritmů jednak jako celku, ale zejména také porovnávání a případná kombinace jednotlivých bloků napříč spektrem kompresních algoritmů.

V dalších kapitolách následuje popis jednotlivých bloků navrženého systému. O každém bloku je nejprve pojednáno obecně, poté je popsána funkce deklarující daný blok. Názvy funkcí jsou pro přehlednost započaty velikým písmenem K, následuje klíčové slovo charakterizující daný systémový blok. Název deklarační funkce končí velikým písmenem X, které je při implementaci nahrazeno klíčovým slovem charakterizujícím implementovaný kompresní algoritmus. Deklarace funkce je provedena pomocí syntaxe programu Matlab. V popisu každé funkce jsou rozebrány všechny vstupní a výstupní parametry.

### 2.1 Předzpracování

Blok nazvaný *Předzpracování* je volitelným blokem. Do systému byl vložen pro zvýšení jeho univerzálnosti. Některé kompresní algoritmy totiž před *Časově-frekvenční analýzou* provádějí



Obrázek 1: Blokové schéma popisovaného experimentálního zvukového kodéru.

úpravu vstupního signálu. Z hlediska kompatibility by se měl formát vstupujících i vystupujících dat shodovat.

Blok je implementován ve formě funkce  
`function [vystup_data, info_pre] = KpreprocX(vstup_data) ,`

kde `vstup_data` je vektor vstupních dat, `vystup_data` vektor výstupních dat a pomocí `info_pre` funkce předává parametry bloku *Sběr vedlejších a pomocných informací*.

## 2.2 Časově-frekvenční analýza

Blok nazvaný *Časově-frekvenční analýza* představuje transformaci vstupního signálu do formy vhodné pro následnou kompresi. V nejjednodušší formě se může jednat například o banku filtrů (například MPEG 1, vrstva 1 a 2), u pokročilejších algoritmů se může jednat třeba o transformaci DCT (například MPEG 1, vrstva 3) a podobně.

Deklarace funkce:

`function [vystup_data,popis_data,info_tf,varargout] = KtimefreqX(vstup_data) ,`

kde `vstup_data` je vektor vstupních dat, `vystup_data` vektor výstupních dat, vektor `popis_data` popisuje strukturu vektoru výstupních dat a pomocí `info_tf` funkce předává parametry bloku *Sběr vedlejších a pomocných informací*. Poslední výstupní parametr `varargout` (variabilní počet dalších výstupních argumentů) je přidán pro zvýšení univerzálnosti systému. Pomocí tohoto parametru se přenášejí informace specifické pro některé pokročilé kompresní algoritmy.

## 2.3 Psychoakustický model

Úkolem *Psychoakustického modelu* je analyzovat vstupní signál z hlediska jeho vnímání lidským sluchem. Výstupem psychoakustického modelu je vždy prahová (maskovací) křivka. Zvuky ležící pod touto křivkou nejsou lidským sluchem vnímány, z tohoto pohledu nenesou žádnou užitečnou informaci. Úkolem dalších systémových bloků je nadbytečnou informaci zanedbat a dále přenášet

pouze údaje popisující složky signálu ležící nad prahovou křivkou.

Deklarace funkce:

```
function [prah_f, prah_spl, info_pm, varargout] = KpsychomodX(vstup_data) ,
```

kde `vstup_data` je vektor vstupních dat, vektory `prah_f` a `prah_spl` popisují výstupní prahovou křivku. Oba vektory mají stejnou velikost, první z parametrů udává frekvence, pro které je vypočítána prahová křivka, druhý pak udává vypočítanou hladinu akustického tlaku (SPL) pro danou frekvenci. Pomocí `info_pm` může funkce předávat parametry bloku *Sběr vedlejších a pomocných informací*. Tento parametr se spolu s posledním výstupním parametrem `varargout` v současnosti nevyužívají, oba byly přidány pro budoucí kompatibilitu.

## 2.4 Alokace bitů a Kvantizér

Bloky *Alokace bitů* a *Kvantizér* se v literatuře v závislosti na autorovi slučují či rozdělují. Alokace bitů a kvantizér spolu každopádně úzce souvisí. Blok *Alokace bitů* má za úkol podle výstupu psychoakustického modelu navrhnout použitou kvantizaci. Pro algoritmy využívající banku filtrů se jedná o počty bitů použité pro kvantizaci signálů v jednotlivých pásmech.

Blok *Kvantizér* realizuje kvantizaci signálu podle výstupu bloku *Alokace bitů*.

Deklarace funkce:

```
function [vystup_data, popis_vystup, info_abk, varargout] = ..  
    KalokqantX(vstup_data, popis_vstup, varargin) ,
```

kde `vstup_data` je vektor vstupních dat, vektor `popis_vstup` popisuje formát vstupních dat. Pro budoucí kompatibilitu a zvýšení univerzálnosti je přidán vstupní parametr `varargin`. Formát dat ve vektoru výstupních dat `vystup_data` popisuje vektor `popis_vystup`, pomocí `info_abk` funkce opět předává parametry bloku *Sběr vedlejších a pomocných informací*. Parametr `varargout` byl podobně jako u ostatních bloků přidán pro budoucí kompatibilitu a větší univerzálnost navrženého systému.

## 2.5 Entropický kodér

Součástí všech prakticky použitelných kodérů zvuku je *Entropický kodér*. Jeho podstatou je bezztrátový kopresní algoritmus, který většinou využívá pravděpodobnosti výskytu jednotlivých symbolů. Podle této pravděpodobnosti přiřazuje často se vyskytujícím symbolům kratší kódové slovo a naopak méně často se vyskytujícím slovům přiřazuje kódová slova delší. Výsledkem takového překódování je snížení redundance. Pokud jsou pravděpodobnosti výskytu jednotlivých vstupních symbolů značně variabilní (s velkým rozptylem), bude datová úspora získaná entropickým kódováním značná. Oproti tomu v případě podobných pravděpodobností výskytu jednotlivých vstupních symbolů bude datová úspora malá.

Nejčastěji používaným entropickým kódovacím algoritmem je Huffmanovo kódování.

Deklarace funkce:

```
function [vystup_data, popis_vystup, info_ek, varargout] = ..  
    KentropcoderX(vstup_data, popis_vstup, varargin) ,
```

Popis jednotlivých argumentů se shoduje s předchozím blokem *Alokace bitů a Kvantizér*. Parametr `info_ek` předává informace bloku *Sběr vedlejších a pomocných informací*.

## 2.6 Sběr vedlejších a pomocných informací

Blok nazvaný *Sběr vedlejších a pomocných informací* byl do systému začleněn pro jeho zpřehlednění a zjednodušení. Má za úkol shromažďovat informace potřebné pro správné dekodování výstupu. Do tohoto bloku vstupují informace ze všech ostatních bloků systému. Jedná se například o údaje popisující implementovaný algoritmus. Pro systém MP3 by se jednalo například o údaj o kvantizaci, počtu činitelů měřítka a jejich velikosti nebo o délku bloků.

Deklarace funkce:

```
function [info, popis_info] = KsideinfoX(varargin) ,
```

Do bloku *Sběr vedlejších a pomocných informací* vstupují data ze všech ostatních bloků přes argument `varargin`. Tento parametr je použit pro zpřehlednění, nahrazuje parametry `info_pre`, `info_tf`, `info_pm`, `info_abk` a `info_ek`. Díky deklaraci pomocí `varargin` je navíc zajištěna budoucí kompatibilita, neboť je možné přidávat další vstupní parametry. Struktura výstupního parametru `info` je popsána pomocí druhého výstupního parametru `popis_info`.

## 2.7 Formátování výstupního datového toku

Výstupní datový tok, respektive soubor odpovídající danému implementovanému algoritmu, vytváří blok nazvaný *Formátování výstupního datového toku*. Jeho úkolem je vzít surová zkomprimovaná data a všechny pomocné informace, z nich vytvořit jednotlivé rámce, vytvořit výstupní bitový tok a ten následně uložit do odpovídajícího výstupního souboru.

Deklarace funkce:

```
function [status] = KformoutX(info, popis_info, vstup_data, popis_vstup, varargin) ,
```

Do bloku vstupují jednak pomocná a vedlejší data z bloku *Sběr vedlejších a pomocných informací* pomocí parametrů `info` a `popis_info`, dále pak data z bloku *Entropický kodér* přes parametry `vstup_data`, `popis_vstup` a `varargin`. Funkce zpracuje všechna vstupní data a vrací pouze informativní hodnotu `status`. Je-li rovna nule, zakódování proběhlo bez problémů.

## 3 Závěr

V článku byl prezentován návrh univerzálního modulárního systému pro kompresi zvuku. Navržený systém umožňuje implementaci současných i nových perceptuálních kodérů za účelem jejich porovnání, optimalizace a vylepšení. Zejména při vývoji nových způsobů komprese zvukových signálů umožňuje navržený kodér spolupráci týmu vývojářů a příslušné rozdělení práce. Implementace je připravena v prostředí Matlab, které je rozšířené v prostředí technických univerzit a tudíž velmi vhodné pro tento projekt.

## Poděkování

Popisovaný výzkum je podporován grantem České grantové agentury číslo 102/05/2054 – Kvalitativní aspekty zpracování audiovizuální informace v multimediálních systémech.

## Reference

- [1] Bosi, M., Goldberg, R. E.: *Introduction to Digital Audio Coding and Standards*, Kluwer Academic Publishers, USA 2003.

- [2] Painter, T., Spanias, A.: *Perceptual Coding of Digital Audio*, In IEEE Proceedings, Vol 88, No. 4, pages 451-513, April 2000
- 

František Rund

Katedra radioelektroniky, FEL, ČVUT Praha, Technická 2, 166 27, Praha 6  
tel. 22435 2108, e-mail: xrund@fel.cvut.cz

Jan Nováček

Katedra radioelektroniky, FEL, ČVUT Praha, Technická 2, 166 27, Praha 6  
e-mail: novacj1@fel.cvut.cz