

# Dialogové systémy

Luděk Bártek

Laboratoř vyhledávání a dialogu, Fakulta Informatiky Masarykovy Univerzity,  
Brno

jaro 2013

- Zkoumá zvukovou stránku jazyka z různých aspektů.
- Základní pojmy, které souvisejí s dialogovými systémy:
  - foném
    - samohlásky – formanty
    - souhlásky – znělost/neznělost souhlásek
  - koartikulace
  - spodoba znělosti

# Fonémy a fonetická transkripce

Dialogové  
systémy

Luděk Bártek

Základy  
fonetiky

Úvod do  
počítačového  
zpracování  
zvuku

Komunikace  
uživatel –  
dialogový  
systém

VoIP  
SIP

Zpracování di-  
gitalizovaného  
signálu

Analýza  
digitalizovaného  
signálu v časové  
oblasti

Analýza  
digitalizovaného  
signálu ve  
frekvenční  
oblasti

- Foném – elementární zvukový segment, který je vymezen na základě své schopnosti diferencovat vyšší, znakové jednotky jazykového systému (morfémy).
- Fonetická transkripce (přepis) – převod psaného textu do odpovídající fonetické podoby:

na shledanou → na zhledanou | na schledanou

- Fonetická abeceda – slouží k zápisu fonetického přepisu
  - Mezinárodní fonetická abeceda (IPA) – součástí standardu UNICODE
  - Řečové vyhodnocení metod fonetické abecedy (SAMPa) – sedmibitový přepis fonetické abecedy, využívá se při automatizovaném zpracování (např. řečový syntetizér MBrola, ... )..

- Samohláska – samostatně tvoří slabiku
- Rozdělení samohlásek:
  - krátké: a, e, i, o, u
  - dlouhé: á, é, í, ó, ú
  - dvojhlásky: eu, au, ou
- Obsahují:
  - základní hlasivkový tón – frekvence kmitání hlasivek (100 — 400 Hz)
  - formanty – frekvence vzniklé a zesílené rezonancí v hlasových dutinách.

- Frekvence vzniklé a zesílené rezonancí v hlasových dutinách
  - F1 – vzniká rezonancí v dutině ústní.
  - F2 – vzniká rezonancí v dutině hrdelní.
- Existují i vyšší formanty (F3, ...) – výskyt je často individuální.
- Výskyt a intenzita formantů se může lišit v závislosti na:
  - pohlaví – muž/žena
  - věku – dětství/dospívání/dospělost/seniorský věk
  - zdravotním stavu – např. nachlazení, ochraptělost, nemoci hlasivek a hrtanu, ...
  - ...

# Formanty F1 a F2 pro české samohlásky

Dialogové  
systémy

Luděk Bártek

Základy  
fonetiky

Úvod do  
počítačového  
zpracování  
zvuku

Komunikace  
uživatel –  
dialogový  
systém

VoIP  
SIP

Zpracování di-  
gitalizovaného  
signálu

Analýza  
digitalizovaného  
signálu v časové  
oblasti

Analýza  
digitalizovaného  
signálu ve  
frekvenční  
oblasti

Samohláska	Formant F1	Formant F2
a	700 — 1100 Hz	1100 — 1500 Hz
e	500 — 700 Hz	1500 — 2000 Hz
i	300 — 500 Hz	2000 — 3000 Hz
o	500 — 700 Hz	900 — 1200 Hz
u	300 — 500 Hz	600 — 1000 Hz

Tabulka : Formanty F1 a F2 u samohlásek

# Četnost výskytu samohlásek

Dialogové  
systémy

Luděk Bártek

Základy  
fonetiky

Úvod do  
počítačového  
zpracování  
zvuku

Komunikace  
uživatel –  
dialogový  
systém

VoIP  
SIP

Zpracování di-  
gitalizovaného  
signálu

Analýza  
digitalizovaného  
signálu v časové  
oblasti

Analýza  
digitalizovaného  
signálu ve  
frekvenční  
oblasti

Samohláska(y)	Relativní četnost
[e]	10 %
[a], [o], [i]	6 — 7 %
[í]	4 %
[á], [u], [é], [ou], [ú]	< 4 %
[ó], [au], [eu]	pouze nepatrná frekvence

- Na rozdíl od samohlásek jsou souhlásky dynamické děje.
- Silně závisí na kontextu, ve kterém se nacházejí.
- Tónový charakter mají pouze části některých souhlásek:
- Dělí se podle:
  - znělé – vznikají v hltanu, obsahují základní hlasivkový tón.
  - neznělé – vznikají v řečových dutinách (nosohltanové, ústní, ...), mohou mít charakter šumu (např. sykavky):
    - problematická detekce začátku promluvy při zašuměném zdroji.
  - Znělé a neznělé samohlásky se mohou vyskytovat v párech (párové souhlásky) např.:
    - r/l
    - b/p
    - d/t
    - ...



- Kroky digitalizace zvuku:
  - 1 vzorkování – snímání aktuální hodnoty signálu s danou frekvencí (vzorkovací frekvence)
  - 2 kvantizace – převod reálných hodnot na celočíselné
  - 3 kódování průběhu vlny – způsob ukládání informací o průběhu zvuku.

- Snímání aktuální hodnoty signálu s danou frekvencí – vzorkovací frekvence.
- Vzorkovací frekvence – měla by být minimálně dvojnásobkem nejvyšší frekvence, která je v signálu přítomna, aby bylo možné původní signál bez ztráty informace zrekonstruovat (Shannonův vzorkovací teorém).
- Získané hodnoty musí být následně kvantizovány a vhodným způsobem uloženy.
- Nejpoužívanější vzorkovací frekvence:
  - 8 kHz – telefonní kvalita
  - 16 kHz
  - 22050 Hz – rozhlasová kvalita
  - 44100 Hz – CD kvalita
  - 48 kHz – DVD kvalita

- Metoda převodu spojitych hodnot na diskretní.
- Princip:
  - Pokud hodnota signálu překročí  $n$ . násobek kvantizačního kroku je jí přiřazena hodnota  $n$ .
  - kvantizační krok = rozsah hodnot měřené veličiny/počet diskretních hodnot
  - kvantizační chyba – zaokrouhlovací chyba způsobená velikostí kvantizačního kroku, přímo úměrná velikosti kvantizačního kroku.
- Běžně používané kvantizace:
  - zpracování zvuku:
    - $2^8$
    - $2^{16}$
    - $2^{24}$
  - zpracování obrazu, ... navíc
    - $2^{32}$

# Způsoby kódování průběhu vlny

Dialogové  
systémy

Luděk Bártek

Základy  
fonetiky

Úvod do  
počítačového  
zpracování  
zvuku

Komunikace  
uživatel –  
dialogový  
systém

VoIP  
SIP

Zpracování di-  
gitalizovaného  
signálu

Analýza  
digitalizovaného  
signálu v časové  
oblasti

Analýza  
digitalizovaného  
signálu ve  
frekvenční  
oblasti

- Přímé ukládání hodnot získaných kvantizací – kódování PCM (Pulse-Code Modulation).
  - relativně pomalé změny průběhu zvukového signálu – malé rozdíly mezi sousedními vzorky.
  - Velká redundance dat.
  - Problém v případě příliš velkého rozptylu amplitud v signálu – příliš velký kvantizační krok – příliš velká kvantizační chyba, příliš malý kvantizační krok – přetečení v okamžiku zvětšení amplitudy signálu.
- Diferenční PCM – ukládá se rozdíl mezi sousedními vzorky
- Adaptivní diferenční PCM — diferenční PCM s proměnou velikostí kvantizačního kroku.

# Diferenční pulsní kódová modulace

Dialogové systémy

Luděk Bártek

Základy fonetiky

Úvod do počítačového zpracování zvuku

Komunikace uživatel – dialogový systém

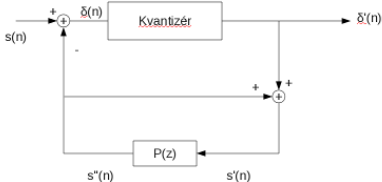
VoIP  
SIP

Zpracování digitalizovaného signálu

Analýza digitalizovaného signálu v časové oblasti

Analýza digitalizovaného signálu ve frekvenční oblasti

- Vychází z předpokladů:
  - Rozdíl dvou po sobě jdoucích vzorků je podstatně menší hodnota než hodnota vzorku.
  - Následující vzorek lze poměrně přesně odhadnout jako lineární kombinaci předchozích vzorků.
- Blokové schéma kódování signálu pomocí DPCM



- $s''(n)$  – odhad hodnoty řečového vzorku
- $s'(n)$  – rekonstruovaný signál, získaný jako součet kvantizovaného signálu  $\delta'(n)$  a  $s''(n)$
- $\delta(n) = s(n) - s''(n)$

# Adaptivní pulsní kódová modulace

Dialogové  
systémy

Luděk Bártek

Základy  
fonetiky

Úvod do  
počítačového  
zpracování  
zvuku

Komunikace  
uživatel –  
dialogový  
systém  
VoIP  
SIP

Zpracování di-  
gitalizovaného  
signálu

Analýza  
digitalizovaného  
signálu v časové  
oblasti

Analýza  
digitalizovaného  
signálu ve  
frekvenční  
oblasti

- Možné velké změny amplitudy signálu:
  - Nepřesné zachycení slabého signálu – amplituda je příliš malá, srovnatelná s kvantizačním krokem (příliš velký kvantizační krok).
  - Zkreslení (ořezání) silného signálu – dojde k přetečení rozsahu hodnot určených pro zakódování signálu (příliš malý kvantizační krok).
- Řešení: přizpůsobení kvantizačního kroku amplitudě signálu.

# Způsoby komunikace uživatele s dialogovým systémem

Dialogové systémy

Luděk Bártek

Základy fonetiky

Úvod do počítačového zpracování zvuku

Komunikace uživatel – dialogový systém

VoIP  
SIP

Zpracování digitalizovaného signálu

Analýza digitalizovaného signálu v časové oblasti

Analýza digitalizovaného signálu ve frekvenční oblasti

- Hlasová:
  - komunikace většinou prostřednictvím telefonní sítě (PSTN, VoIP).
  - Digitalizace hlasu probíhá:
    - Na straně uživatele – komunikace pomocí VoIP.
    - Na straně telefonní ústředny – DS používá VoIP, uživatel používá PSTN.
    - Na straně DS – uživatel i DS používají PSTN.
  - Rozpoznávání řeči probíhá většinou na straně DS.

# Způsoby komunikace uživatele s dialogovým systémem

Dialogové systémy

Luděk Bártek

Základy fonetiky

Úvod do počítačového zpracování zvuku

Komunikace uživatel – dialogový systém

VoIP  
SIP

Zpracování digitalizovaného signálu

Analýza digitalizovaného signálu v časové oblasti

Analýza digitalizovaného signálu ve frekvenční oblasti

- textová:
  - uživatel komunikuje s DS buď pomocí specializovaného klienta nebo pomocí běžných protokolů z rodiny TCP/IP.
  - Odpadá nutnost rozpoznávání řeči.
  - Využívá se hlavně pro vývoj a ladění.
- hlasová+textová:
  - komunikace s DS buď VoIP nebo specializovaný klient.
  - V případě VoIP text buď pomocí DTMF nebo simulace SMS.



- VoIP – rodina protokolů pro řízení průběhu hlasové komunikace a přenos hlasu přes internet (sít' na bázi IP).
- Využívá se pro IP telefonii.
- Využívá protokoly:
  - UDP (transportní vrstva):
    - Stará se o přenos paketů přes počítačovou sít' mezi dvěma body.
    - Není zajištěno doručení paketů ani jejich pořadí.
    - Výhoda – nízká rezie přenosu dat.
    - Nevýhody – možná ztráta dat a možnost velkých rozdílů v rychlosti doručení jednotlivých paketů
  - RTP (relační vrstva):
    - Využívá se pro přenos multimediálních dat.
    - Zajišťuje doručení paketů.
    - Umožňuje řízení parametrů přenosu – zajistí malé rozdíly v rychlosti doručení paketů.

- VoIP – řada implementací
- liší se
  - použitými standardy – H.323 (na ústupu, standard ITU, komplexní, relativně komplikovaný), SIP (jednodušší náhrada H.323, v současnosti velmi rozšířený), firemní – Skinny (Cisco), HFA (Siemens), ...
  - službami – telefonie, TV (DVB), fax, zasílání zpráv, ...
  - signalizací – závisí na zvoleném standardu a použitých protokolech.
  - ...

# Session Initiation Protocol (SIP)

Dialogové  
systémy

Luděk Bártek

Základy  
fonetiky

Úvod do  
počítačového  
zpracování  
zvuku

Komunikace  
uživatel –  
dialogový  
systém

VoIP  
SIP

Zpracování di-  
gitalizovaného  
signálu

Analýza  
digitalizovaného  
signálu v časové  
oblasti

Analýza  
digitalizovaného  
signálu ve  
frekvenční  
oblasti

- protokol pro řízení signalizace pro VoIP na aplikační vrstvě OSI modelu
- textový protokol pracující v režimu klient–server, poskytující mechanismy pro:
  - přesměrování hovoru
  - číselnou identifikaci volajícího a volaného
  - osobní mobilitu
  - autentizaci volajícího a volaného
  - podporu konferenčních hovorů prostřednictvím vícesměrového zasílání dat (multicast).
  - ...

# SIP – pokračování

Dialogové  
systémy

Luděk Bártek

Základy  
fonetiky

Úvod do  
počítačového  
zpracování  
zvuku

Komunikace  
uživatel –  
dialogový  
systém

VoIP  
SIP

Zpracování di-  
gitalizovaného  
signálu

Analýza  
digitalizovaného  
signálu v časové  
oblasti

Analýza  
digitalizovaného  
signálu ve  
frekvenční  
oblasti

- Identifikace účastníka – URI ve tvaru *sip:číslo@adresa\_počítače*
  - číslo – číslo přidělené uživateli na daném stroji (VoIP ústředně)
  - adresa počítače – adresa (FQDN/IP) ústředny, na které je uživatel registrován.
- SIP relace může být:
  - přímá – navázána přímo komunikujícími stranami
  - s použitím SIP proxy serveru/ů – tyto slouží jako registrátoři účastníků.
- Činnosti protokolu SIP:
  - Lokalizace účastníka – pomocí identifikace
  - Zjištění stavu účastníka – připravenost k přijetí hovoru vs. obsazeno/přesměrováno
  - Zjištění možností účastníka – dostupné kodeky, dostupná šířka pásma, podpora audia/video, ...
  - Vlastní navázání spojení – využívá se protokol SDP
    - popisuje navazované spojení,
    - odkazuje na RTP datový tok, který je využit pro

# Řízení průběhu spojení pomocí protokolu SIP

Dialogové systémy

Luděk Bártek

Základy fonetiky

Úvod do počítačového zpracování zvuku

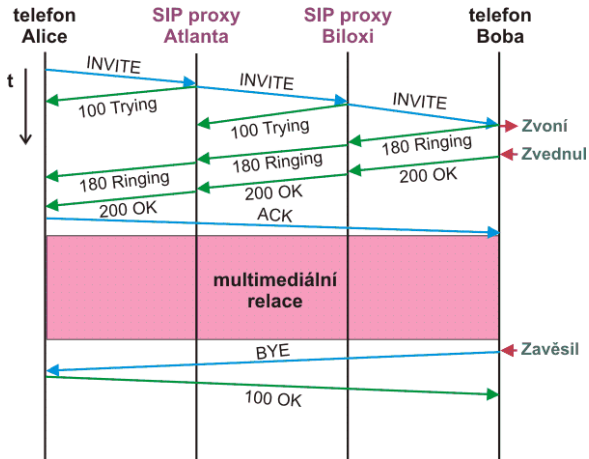
Komunikace uživatel – dialogový systém

VoIP  
SIP

Zpracování digitalizovaného signálu

Analýza digitalizovaného signálu v časové oblasti

Analýza digitalizovaného signálu ve frekvenční oblasti



# Zpracování digitalizovaného signálu

Dialogové  
systémy

Luděk Bártek

Základy  
fonetiky

Úvod do  
počítačového  
zpracování  
zvuku

Komunikace  
uživatel –  
dialogový  
systém

VoIP  
SIP

Zpracování di-  
gitalizovaného  
signálu

Analýza  
digitalizovaného  
signálu v časové  
oblasti

Analýza  
digitalizovaného  
signálu ve  
frekvenční  
oblasti

- Zvuk je neměnný pouze na krátkých časových úsecích – metody krátkodobé analýzy.
- Tento interval se nazývá mikrosegment – velikost 10 — 40 ms.
- Metody krátkodobé analýzy:
  - V časové oblasti – zpracovávají se přímo hodnoty jednotlivých vzorků.
  - Ve frekvenční oblasti – ze vzorků se získávají frekvenční charakteristiky, které jsou následně zpracovány.
- Modelování funkce Cortiho ústrojí – pomocí diferenciálních rovnic se simuluje rezonance na určitých vláčkách Cortiho ústrojí.

# Váhové okénko

Dialogové  
systémy

Luděk Bártek

Základy  
fonetiky

Úvod do  
počítačového  
zpracování  
zvuku

Komunikace  
uživatel –  
dialogový  
systém  
VoIP  
SIP

Zpracování di-  
gitalizovaného  
signálu

Analýza  
digitalizovaného  
signálu v časové  
oblasti

Analýza  
digitalizovaného  
signálu ve  
frekvenční  
oblasti

- Při krátkodobé analýze předpokládáme, že signál je v okolí mikrosegmentu periodický se stejnou periodou jako uvnitř.
- Vzniklá chyba se kompenzuje použitím „okénka”.
- Okénko – posloupnost vah pro vzorky v mikrosegmentu.
- Tyto váhy by měly odpovídat tomu, jak je daný vzorek ovlivněn okolím mikrosegmentu.
- Nejčastěji používané typy okének:
  - pravoúhlé okénko
  - Hammingovo okénko

# Hammingovo okénko

Dialogové  
systémy

Luděk Bártek

Základy  
fonetiky

Úvod do  
počítačového  
zpracování  
zvuku

Komunikace  
uživatel –  
dialogový  
systém

VoIP  
SIP

Zpracování di-  
gitalizovaného  
signálu

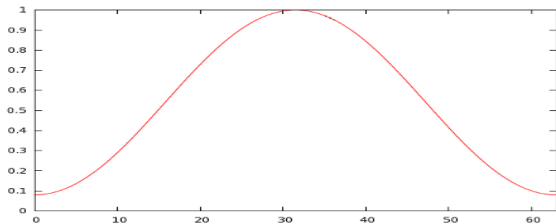
Analýza  
digitalizovaného  
signálu v časové  
oblasti

Analýza  
digitalizovaného  
signálu ve  
frekvenční  
oblasti

- Vychází z předpokladu, že čím jsou vzorky blíže středu mikrosegmentu, tím méně jsou ovlivněny okolím.
- Pro výpočet vah se používá vzorec:

$$w(n) = \begin{cases} n = 0 \dots N - 1 & 0.54 - 0.46 \cos\left(\frac{2\pi n}{N-1}\right) \\ n < 0 \vee n \geq N & 0 \end{cases}$$

- Průběh vah okénka na mikrosegmentu:





- Vychází se z předpokladu:
  - 1 vzorky mikrosegmentu nejsou pro naše potřeby ovlivněny okolím mikrosegmentu
  - 2 všechny vzorky mikrosegmentu jsou ovlivněny stejně.
- Všechny vzorky mikrosegmentu mají shodnou váhu.

$$w(n) = \begin{cases} 0 \leq n < N & 1 \\ n < 0 \vee n \geq N & 0 \end{cases}$$

# Analýza digitalizovaného signálu v časové oblasti

Dialogové  
systémy

Luděk Bártek

Základy  
fonetiky

Úvod do  
počítačového  
zpracování  
zvuku

Komunikace  
uživatel –  
dialogový  
systém

VoIP  
SIP

Zpracování di-  
gitalizovaného  
signálu

Analýza  
digitalizovaného  
signálu v časové  
oblasti

Analýza  
digitalizovaného  
signálu ve  
frekvenční  
oblasti

- Vychází přímo z hodnot vzorků, nikoliv z hodnot spektra.
- Používané metody:
  - funkce krátkodobé energie
  - funkce krátkodobé intenzity
  - funkce středního počtu průchodů nulou
  - diference 1. řádu
  - autokorelační funkce
  - ...

# Analýza v časové oblasti

## Funkce krátkodobé energie

Dialogové  
systémy

Luděk Bártek

Základy  
fonetiky

Úvod do  
počítačového  
zpracování  
zvuku

Komunikace  
uživatel –  
dialogový  
systém  
VoIP  
SIP

Zpracování di-  
gitalizovaného  
signálu

Analýza  
digitalizovaného  
signálu v časové  
oblasti

Analýza  
digitalizovaného  
signálu ve  
frekvenční  
oblasti

- Využívá funkci průměrné energie v rámci segmentu:

$$E(n) = \sum_{k=-\infty}^{\infty} (s(k)\omega(n-k))^2$$

- $s(k)$  – vzorek v čase  $k$
- $\omega(n-k)$  – váha odpovídajícího okénka pro čas  $k$
- Výstupem je průměrná energie v daném okénku.
- Druhá mocnina zvyšuje dynamiku zvukového signálu.
- Použití:
  - automatické oddělení ticha řeči (signálu)
  - příznaky v jednoduchých klasifikátorech slov
  - oddělení znělých a neznělých částí promluvy.

- Funkce intenzity signálu v daném okénku.

$$I(n) = \sum_{k=-\infty}^{\infty} |s(k)|\omega(n - k)$$

- $|s(k)|$  – absolutní hodnota vzorku v čase  $k$
- $\omega(n - k)$  – váha odpovídajícího okénka pro čas  $k$
- Použití – stejné jako funkce krátkodobé energie.
- Oproti krátkodobé energii nezvýrazňuje tolik dynamiku řečového signálu.

# Analýza v časové oblasti

Krátkodobá funkce středního počtu průchodu nulou

Dialogové  
systémy

Luděk Bártek

Základy  
fonetiky

Úvod do  
počítačového  
zpracování  
zvuku

Komunikace  
uživatel –  
dialogový  
systém  
VoIP  
SIP

Zpracování di-  
gitalizovaného  
signálu

Analýza  
digitalizovaného  
signálu v časové  
oblasti

Analýza  
digitalizovaného  
signálu ve  
frekvenční  
oblasti

- Počítá změny znaménka digitalizovaného signálu.

$$Z(n) = \sum_{k=-\infty}^{\infty} |\operatorname{sgn}[s(k)] - \operatorname{sgn}[s(k-1)]| \omega(n-k)$$

- Varianta – počet lokálních extrémů.
- Obě metody mohou být negativně zatíženy šumem zvukového pozadí.
- Použití:
  - detekce ticha
  - detekce začátku a konce i zašuměné promluvy
  - přibližné určení základního hlasivkového tónu a formantů
  - příznaky jednodušších klasifikátorů slov

# Analýza v časové oblasti

## Autokorelační funkce

Dialogové  
systémy

Luděk Bártek

Základy  
fonetiky

Úvod do  
počítačového  
zpracování  
zvuku

Komunikace  
uživatel –  
dialogový  
systém

VoIP  
SIP

Zpracování di-  
gitalizovaného  
signálu

Analýza  
digitalizovaného  
signálu v časové  
oblasti

Analýza  
digitalizovaného  
signálu ve  
frekvenční  
oblasti

- Vrací podobnost úseků daného mikrosegmentu (čím větší výsledná hodnota, tím podobnější úseky posunuté o  $m$  vzorků).

$$R(m, n) = \sum_{k=-\infty}^{\infty} (s(k)\omega(n-k))(s(k+m)\omega(n-k+m))$$

- Je-li signál periodický s periodou  $P$ ,  $R(m, n)$  nabývá maxima pro  $m=0, P, 2P, \dots$
- Předpokládá délku mikrosegmentu aspoň  $2P$ .
- Použití:
  - Používá se k zjišťování periodicity signálu základního tónu řeči.
  - Základ pro výpočet koeficientů LPA

# Analýza signálu ve frekvenční oblasti

Dialogové  
systémy

Luděk Bártek

Základy  
fonetiky

Úvod do  
počítačového  
zpracování  
zvuku

Komunikace  
uživatel –  
dialogový  
systém

VoIP  
SIP

Zpracování di-  
gitalizovaného  
signálu

Analýza  
digitalizovaného  
signálu v časové  
oblasti

Analýza  
digitalizovaného  
signálu ve  
frekvenční  
oblasti

- Transformuje digitální řečový signál z časové oblasti do frekvenční oblasti.
- Využívá k tomu nejčastěji Fourierovu transformaci.
- Nejčastěji používané druhy analýzy ve frekvenční oblasti:
  - krátkodobá Fourierova transformace
  - krátkodobá diskrétní Fourierova transformace
  - rychlá Fourierova transformace
  - kepstrální analýza
  - lineární predikce
  - ...

# Analýza signálu ve frekvenční oblasti

## Krátkodobá Fourierova transformace

Dialogové  
systémy

Luděk Bártek

Základy  
fonetiky

Úvod do  
počítačového  
zpracování  
zvuku

Komunikace  
uživatel –  
dialogový  
systém  
VoIP  
SIP

Zpracování di-  
gitalizovaného  
signálu

Analýza  
digitalizovaného  
signálu v časové  
oblasti

Analýza  
digitalizovaného  
signálu ve  
frekvenční  
oblasti

- Vychází z Fourierovy transformace:

$$S(\omega, t) = \sum_{k=-\infty}^{\infty} s(k)h(t - k)e^{-j\omega k}$$

- Obyčejnou Fourierovu transformaci získáme fixací času  $t$ .
  - $|S(\omega, t)|$  – amplituda složky akustického spektra odpovídající frekvenci  $\omega$  v čase  $t$ .
  - $h(n)$  – váhová funkce okénka.
- Předpokládá na vstupu periodickou funkci – zvuk je periodický na krátkých časových úsecích.
  - Při jejím použití se předpokládá, že zpracováváný mikrosegment se periodicky opakuje.



# Analýza signálu ve frekvenční oblasti

## Diskrétní Fourierova transformace

Dialogové systémy

Luděk Bártek

Základy fonetiky

Úvod do počítačového zpracování zvuku

Komunikace uživatel – dialogový systém

VoIP  
SIP

Zpracování digitalizovaného signálu

Analýza digitalizovaného signálu v časové oblasti

Analýza digitalizovaného signálu ve frekvenční oblasti

- Používá se pro vyjádření spektrálních vlastností periodických posloupností s periodou  $N$  vzorků resp. konečných posloupností délky  $N$  vzorků.
- Výpočet koeficientů  $X(k)$  DFT:

$$X(k) = \sum_{n=0}^{N-1} x(n) \exp(-j \frac{2\pi}{N} kn) = \sum_{n=0}^{N-1} x(n) W_N^{-kn}$$

- $|X(k)|$  – intenzita  $k$ . spektrálního koeficientu, frekvence závisí na velikosti mikrosegmentu  $N$  a vzorkovací frekvenci.
- $x(n)$  –  $n$ . vzorek daného mikrosegmentu
- $W_n = \exp(j * 2\pi / N) = \cos(2\pi / N) + j * \sin(2\pi / N)$ .
- Výpočet  $n$ . vzorku na základě hodnot  $X(k)$  – IDFT:

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) \exp(j \frac{2\pi}{N} kn) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) W_N^{kn},$$

# Analýza signálu ve frekvenční oblasti

Rychlá diskrétní Fourierova transformace

Dialogové  
systémy

Luděk Bártek

Základy  
fonetiky

Úvod do  
počítačového  
zpracování  
zvuku

Komunikace  
uživatel –  
dialogový  
systém

VoIP  
SIP

Zpracování di-  
gitalizovaného  
signálu

Analýza  
digitalizovaného  
signálu v časové  
oblasti

Analýza  
digitalizovaného  
signálu ve  
frekvenční  
oblasti

- Výpočet spektrálních koeficientů pomocí DFT –  $n^2$  operací nad komplexními čísly.
- Pomocí FFT –  $N * \log_2 N / 2$  operací násobení.
- FFT požaduje, aby délka analyzovaného segmentu byla mocninou 2.

# Analýza signálu ve frekvenční oblasti

## Kepstrální analýza

Dialogové  
systémy

Luděk Bártek

Základy  
fonetiky

Úvod do  
počítačového  
zpracování  
zvuku

Komunikace  
uživatel –  
dialogový  
systém

VoIP  
SIP

Zpracování di-  
gitalizovaného  
signálu

Analýza  
digitalizovaného  
signálu v časové  
oblasti

Analýza  
digitalizovaného  
signálu ve  
frekvenční  
oblasti

- Vychází z modelu činnosti hlasového ústrojí.
- Řečové kmity lze modelovat jako odezvu lineárního systému na buzení sestávající ze sledu pulzů pro znělou řeč a šumu pro neznělou.
- Kepstrum –  $X(k) = IFFT(FFT(x(k)))$
- Kepstrální analýza umožňuje z řeči oddělit parametry buzení a parametry hlasového ústrojí.
- Využití:
  - ocenění fonetické struktury řeči – znělost perioda základního tónu, formanty, ...
  - rozpoznávání slov
  - verifikace a identifikace mluvčího
  - ...

# Analýza signálu ve frekvenční oblasti

## Lineární prediktivní analýza

Dialogové  
systémy

Luděk Bártek

Základy  
fonetiky

Úvod do  
počítačového  
zpracování  
zvuku

Komunikace  
uživatel –  
dialogový  
systém

VoIP  
SIP

Zpracování di-  
gitalizovaného  
signálu

Analýza  
digitalizovaného  
signálu v časové  
oblasti

Analýza  
digitalizovaného  
signálu ve  
frekvenční  
oblasti

- Jedna z nejefektivnějších metod analýzy akustického signálu – zajišťuje velmi přesné odhady parametrů při relativně malé zátěži.
- Vychází z předpokladu, že  $s(k)$  lze popsat jako lineární kombinaci  $N$  předchozích vzorků a buzení  $u(k)$ :

$$s(k) = - \sum_{i=1}^N a_i s(k-i) + Gu(k)$$

kde  $G$  je koeficient zesílení a  $N$  řád modelu.

- Použití:
  - určování spektrálních charakteristik modelu hlasového ústrojí
  - z chyby predikce lze odvodit poznatky o znělosti a určit frekvenci základního hlasivkového tónu
  - koeficienty  $a_i$  nesou informaci o spektrálních vlastnostech – lze je použít jako příznaky pro rozpoznávání řeči.