

PB095 - Úvod do počítačového zpracování řeči

Luděk Bártek¹

1

Obsah

- 1 Organizace předmětu
- 2 Stručná historie zpracování zvuku a řeči
- 3 Akustika

Cíle předmětu

Seznámení s oblastmi:

- digitálního zpracování zvuku
 - v časové oblasti
 - ve frekvenční oblasti
 - převod signálu z časové do frekvenční oblasti
- syntézy řeči
- rozpoznávání řeči
- dialogových systémů

Organizace předmětu

- Dvuhodinová přednáška
- Možnosti zakončení:
 - zkouška - písemka + ústní zkouška, termíny budou vypsány v IS MU během prosince
 - kolokvium - ústní rozprava na dané téma z oblasti zpracování zvuku
 - zápočet

Vyučující

- Luděk Bártek
- e-mail: bar@fi.muni.cz
- kancelář: B314
- konzultace viz osobní stránka v ISu
(<https://is.muni.cz/auth/osoba/2154#vyuka>)

Doporučená literatura

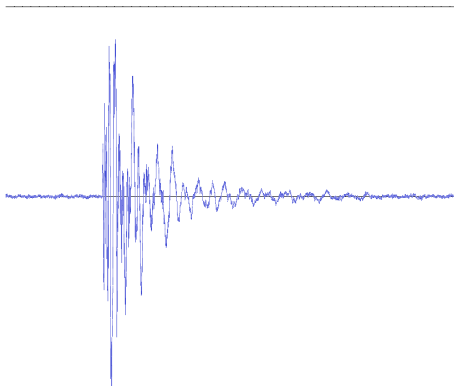
- J. Psutka et al, Mluvíme s počítačem česky, Academia 2006
- J. Psutka, Komunikace s počítačem mluvenou řečí, Academia, Praha, 1995
- Z. Kotek, V. Minařík, Metody rozpoznávání a jejich aplikace, Academia, Praha, 1993
- T. Dutoit, An introduction to Text-to-Speech Synthesis, Kluwer Acad. Publ., 1999
- M. R. Schroeder, Computer Speech, Springer 1999
- Původní stránky předmětu
(<http://www.fi.muni.cz/~kopecek/upzr.htm>) doc. Kopečka
- Stránky Voice Browser Activity (???)

- Laboratoře LSD, NLP
- <http://lsd.fi.muni.cz/>
- <http://nlp.fi.muni.cz/>

Co je to zvuk?

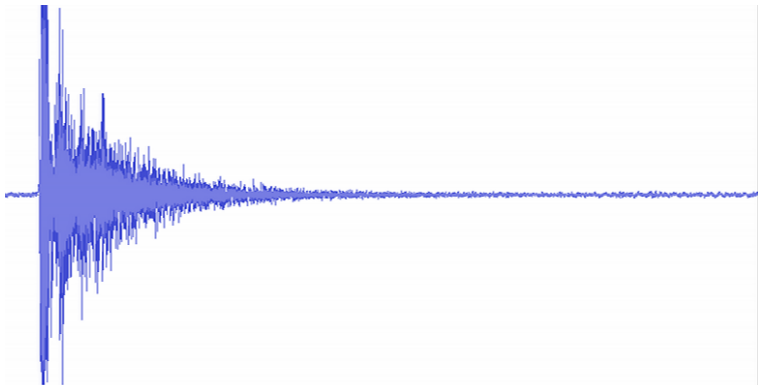
- Akustický signál.
- Jedná se o kmitavý pohyb molekul pružného prostředí.
 - vzduch
 - voda
 - kov
 - ...
- Vyvolán odporem prostředí - vede k opakovanému stlačování prostředí.
- Podrobněji v části fyzikální akustika.

Ukázka zvuku



Zvuk klepnutí na plastové tělo počítače

Ukázka zvuku

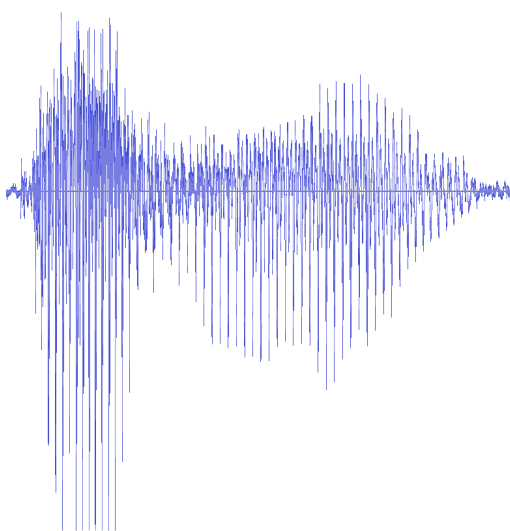


Zvuk tlesknutí

Co je to řeč?

- Akustický signál a gesta sloužící ke komunikaci.
- Obsahuje definované vzory (slova), která jsou dána jazykem.
- Velmi rozvinutý u člověka.
 - Příznaky schopnosti tvorby artikulované řeči již u Australopitéka (-3 milióny let).
 - Slouží ke sdělování: myšlenek, pocitů, emocí, ...
 - myšlenek - "Dnes budeme probírat láčkovce."
 - pocitů - "Je mi krásně.", "Radši se ke mně ani nepřibližuj!", ...
 - emocí - "Au!", "Jé!",
- Určité formy akustické komunikace (řeči) lze pozorovat i u dalších vyšších živočichů:
 - způsob zajištění kooperace při získávání obživy (delfín, vlk, ...)
 - vábení partnera (jelen, ...)
 - vyjádření emočních stavů (pes, opice, ...).
 - ...

Ukázka řeči



Studium řeči

- Fyzika - akustika.
- Biologie - medicína (fyziologie, fyziologická akustika).
- Jazykověda - fonetika.

Přehled historie zpracování a napodobování řeči

- Schopnost artikulované řeči - australopitekus - cca. -3 000 000 let
- Starověk - budování mluvících soch
- Galileo Galilei - souvislost mezi tónem a frekvencí
- 1779 - Christian Gottlieb Kratzenstein - systém rezonátorů pro samohlásky a, e, i, o, u



Přehled historie zpracování a napodobování řeči

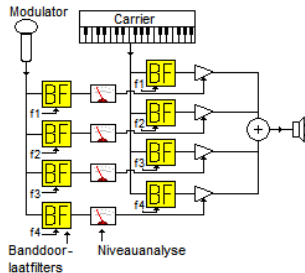
- 1791 - (Johann) Wolfgang von Kempelen (de Pázmánd) - první mechanický řečový syntetizér



- 1835 - zrekonstruován a upraven Wheatonem - navíc pružná "ústní dutina".
- 1846 - J. Faber - mluvící stroj Euphonia

Přehled historie zpracování a napodobování řeči

- 1937 - R. R. Riesz - mechanický mluvící stroj
- 1939 - H. Dudley
 - VODER - elektromechanický řečový syntetizér
 - VOCODER - systém pro kódování a přenos řeči



- 50. léta 20. století - syntéza ve frekvenční oblasti
 - později v časové oblasti
- 70. léta 20. století - počítačové zpracování zvuku

Historie analýzy řeči

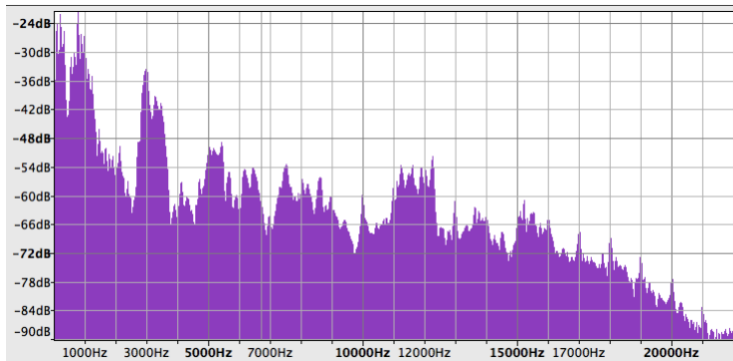
- 19. století porozumění principů tvorby a zpracování řeči (rezonanční teorie, základy fonetiky):
 - J. B. Fourier - Fourierova věta
 - principy spektrální analýzy zvuku
 - H. Helmholtz
 - fyziologie vnímání hudby
 - Helmholtzův rezonátor



Historie analýzy řeči

- Dvacáté století:
 - 1924 - spektrální analýza řeči na bázi formantové analýzy samohlásek
 - Vokodéry - komprese řečového záznamu
 - 1946 - 47 zařízení pro grafické zobrazení řeči
 - 2. polovina 20. století - intenzivní rozvoj teorie a počítačových aplikací

Spektrum zvuku



Textová data k obrázku. (images/spektrum-a.txt)

Současnost

- syntéza řeči:
 - komerční TTS:
 - AT&T Natural Voices
 - IBM Research TTS
 - Loquendo TTS
 - nekomerční TTS:
 - MBrola
 - Festival
 - Demosthenes

Současnost

- Rozpoznávání řeči:
 - izolovaných slov
 - souvislé promluvy
 - komerční: Dragon, ViaVoice Desktop Products
 - nekomerční: Sphinx4, ...
- Dialogové systémy
 - Infocity Liberec (TU Liberec, Prof. J. Nouza), 485353100
 - MIT Cambridge, Spoken Language System Group
 - Mercury - 001-877-648-8255
 - Jupiter - 001-888-573-8255

- Syntéza a rozpoznávání řeči
 - Demosthenes
 - NLP - čeština pro syntetizér MBrola - využit řečový korpus CLAP
- Asistivní technologie:
 - Audi-C - dialogové programování v C++
 - Audis - řečový hypertextový prohlížeč
 - ...
- Dialogové systémy
 - WebGen (<http://lsd.fi.muni.cz/webgen/>) - dialogové generování webových prezentací
 - GATE (<http://lsd.fi.muni.cz/GATE/>) - dialogové kreslení obrázků, dialogové prohlížení obrázků, zvukové zobrazení obrázků
 - ...
- Spolupráce s laboratořemi NLP, VR, ...

Akustika

- Věda zabývající se zvukem.
- Z řeckého akustikos - vztahující se k slyšení.
- Akustika zkoumá zvuk z hlediska:
 - fyzikálního (fyzikální akustika) - zvuk jako fyzikální vlnění
 - rychlost šíření, vztah mezi různými fyzikálními veličinami zvuku, šíření zvuku, ...
 - fyziologického (fyziologická akustika) - tvorba a vnímání řeči u člověka
 - hudebního (hudební akustika) - zvuky a jejich kombinace s ohledem na potřeby hudby
 - jak lidem zní kombinace a sekvence zvuků a tónů, ...
 - molekulárního (molekulární) - vztah molekulární struktury a akustických vlastností
 - k měření se využívá hyperzvuk (≥ 100 MHz).
 - zpracování zvuku na počítači (počítačová akustika) - digitální zpracování zvuku.

- Během semestru se budeme zabývat:
 - fyzikální akustikou
 - fyziologické akustikou
 - počítačovou akustikou