

Dialogové systémy

Luděk Bártek

Laboratoř vyhledávání a dialogu, Fakulta Informatiky Masarykovy Univerzity,
Brno

jaro 2015

Cíl a náplň předmětu

Dialogové
systémy

Luděk Bártek

O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika

- Cíl – seznámení s oblastí dialogových systémů.
- Obsah kurzu:
 - Úvod do dialogových systémů, historie
 - Základní technologie:
 - digitální zpracování zvuku
 - rozpoznávání řeči
 - porozumění rozpoznané řeči
 - syntéza řeči
 - přenos hlasu prostřednictvím počítačové sítě

Cíl a náplň předmětu

pokračování

Dialogové
systémy

Luděk Bártek

O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika

- Obsah kurzu:
 - Dialogové systémy:
 - formální modely dialogu
 - analýza dialogu, kooperativní a nekooperativní dialog
 - dialogové strategie
 - information retrieval DS
 - simulace DS
 - multimodalita
 - nástroje pro tvorbu dialogových systémů – W3C Voice Browser Activity
 - aplikace.

Ukončení předmětu

Dialogové
systémy

Luděk Bártek

O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika

- Možná ukončení:
 - zkouška
 - kolokvium
 - zápočet
- Požadavky:
 - zkouška + kolokvium – dobrá orientace v probírané problematice
 - zkouška – písemka + ústní dozkoušení
 - kolokvium – nástin řešení problémové úlohy a detailnější rozbor některé z použitých technologií.
 - zápočet – zápočtová písemka současně se zkouškou.

Doporučená literatura

Knihy

Dialogové
systémy

Luděk Bártek

O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika

- J. Psutka, Komunikace s počítačem mluvenou řečí, Academia, Praha, 1995
- Z. Kotek, V. Mařík, Metody rozpoznávání a jejich aplikace, Academia, Praha, 1993
- T. Dutoit, An Introduction to Text-to-Speech Synthesis, Kluwer Academic Publishing, 1996
- A. Kobsa, W. Wahlster, User Models in Dialog System, Springer 1989
- D. B. Roe, J. G. Wilpon (editors), Voice Communication Between Humans and Machines, National Academy Press, Washington D.C., 1994
- F. Jelinek, Statistical Methods for Speech Recognition, MIT Press 1997

Doporučená literatura

Web

Dialogové
systémy

Luděk Bártek

O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika

- Nuance Dragon
- Odkazy na dialogové systémy (DS) – Odkazy na dialogové systémy (DS)
- Různé projekty z oblasti počítačové sémantiky na Stanford University
- Stránky W3C Voice Browser Activity
- ...

Co je dialogový systém?

Dialogové
systémy

Luděk Bártek

O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika

- Dialogový systém – systém komunikující s uživatelem pomocí dialogu v přirozeném jazyce
 - většinou se jedná o dialogové rozhraní ke klasickému IS.
- Častá komunikace mluvenou řečí.
- Alternativně:
 - komunikace pomocí DTMF
 - textová komunikace přirozenou řečí
 - multimodální komunikace:
 - řeč + obraz (simulace lidské tváře, titulky ve znakové řeči, ...)
 - řeč + text
 - ...

Výhody a nevýhody dialogových systémů

Dialogové
systémy

Luděk Bártek

O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika

■ Výhody:

- + Přirozenější způsob komunikace.
- + Přístupnost:
 - zrakově a motoricky postižení uživatelé
 - další uživatelé, kterým činí problémy ovládání počítače standardním způsobem
 - možnost podrobnějšího vedení uživatele krok za krokem celým procesem
 - dalším krokem k lepší přístupnosti – multimodální rozhraní.
 - ...

Výhody a nevýhody dialogových systémů

Dialogové
systémy

Luděk Bártek

O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika

■ Výhody:

+ Větší množství potenciálních uživatelů:

- počet uživatelů počítačů a Internetu vs. počet uživatelů telefonu.

+ ...

■ Nevýhody:

- rychlost komunikace

- sekvenční vnímání zvuku vs. paralelní vnímání obrazu
- lze částečně eliminovat pomocí vhodné dialogové strategie

- Laboratoře:
 - LSD – doc. Kopeček
 - zaměřuje se na:
 - dialogové systémy a zpracování zvuku
 - sociální sítě
 - sociální informatika
 - NLP – doc. Pala
 - zaměřuje se na:
 - korpusy
 - slovníky
 - morfologii
 - syntaktickou analýzu
 - sémantiku

Aktuální práce v oblastech souvisejících s dialogovými systémy

Výzkum Česká republika

Dialogové systémy

Luděk Bártek

O předmětu

Úvod do dialogových systémů

Historie zpracování zvuku

Fyzikální akustika

Fyziologická akustika

- FIT VUT Brno
 - analýza signálu
 - rozpoznávání řeči
 - systém pro automatizované zpracování konferencí
 - ...
- ZČU v Plzni
 - rozpoznávání řeči
 - dialogové systémy
 - ...
- ČVUT – syntéza řeči

Aktuální práce v oblastech souvisejících s dialogovými systémy

Komerční sféra - Česká republika

Dialogové systémy

Luděk Bártek

O předmětu

Úvod do dialogových systémů

Historie zpracování zvuku

Fyzikální akustika

Fyziologická akustika

- FROG Systems s.r.o. – CS-voice 97
- OptimSys s.r.o – VoiceXML platforma OptimTalk
- ...

Aktuální práce v oblasti dialogových systémů

Výzkum a práce ve světě

Dialogové
systémy

Luděk Bártek

O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika

- W3C Voice Browser Working Group
 - IBM
 - Nuance Communication
 - Lucent Technologies
 - Motorola
 - ScanSoft
 - Tellme Networks
 - Vocalocity
 - ...
- MIT
- OGI
- EPF Lausane
- ...

European Masters in Language and Speech

Dialogové
systémy

Luděk Bártek

O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika

- Evropské navazující studium v oblastech zpracování řeči a přirozeného jazyka (na FI během magisterského studia).
- Zapojeny VŠ např. v Dánsku, Řecku, Španělsku, Belgii, Německu, Velké Británii, Nizozemí, ...
- Více informací:
 - Stránka o EuroMasters na FI
 - doc. Pala, doc. Kopeček.

Struktura dialogového systému

Dialogové systémy

Luděk Bártek

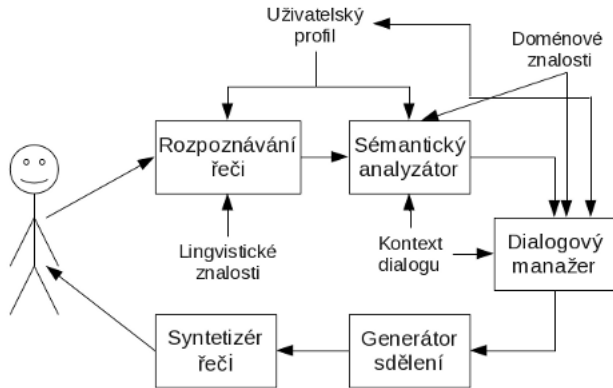
O předmětu

Úvod do dialogových systémů

Historie zpracování zvuku

Fyzikální akustika

Fyziologická akustika



Komponenty dialogového systému

Dialogové
systémy

Luděk Bártek

O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika

- Uživatel – koncové zařízení, které uživateli umožňuje komunikovat s dialogovým systémem:
 - telefon – komunikace prostřednictvím PSTN přes VoIP gateway – VoIP gateway převádí hlas na data a zpět
 - VoIP klient – komunikace prostřednictvím VoIP protokolu přímo s dialogovým systémem (SIP, H.323, Skype, ...)
 - textový klient – komunikace prostřednictvím protokolů DTMF+VoIP protokol, telnet, ssh, XMPP, ...
- Rozpoznávání řeči:
 - převádí mluvené slovo na text
 - využívá se:
 - rozpoznávání plynulé řeči
 - rozpoznávání izolovaných slov
 - pro zvýšení úspěšnosti se používají gramatiky popisující množinu očekávaných vstupů.

Komponenty dialogového systému

Dialogové
systémy

Luděk Bártek

O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika

- Sémantický analyzátor
 - získává relevantní údaje z rozpoznaného textu
 - využívají se např. atributové gramatiky.
- Dialogový manažer
 - konečný automat
 - na základě aktuálního stavu a vstupu od uživatele rozhoduje o dalším průběhu dialogu.
- Generátor promluv – na základě údajů od dialogového manažeru generuje promluvy, které jsou následně syntetizovány.
- Řečový syntetizér – převádí promluvy od generátoru promluv na mluvenou řeč, která je poslána uživateli.

Údaje používané dialogovým systémem

Dialogové
systémy

Luděk Bártek

O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika

- Lingvistické znalosti – údaje o jazyce, které využívá rozpoznávač řeči pro zvýšení úspěšnosti (pravděpodobnosti výskytů jednotlivých sekvencí řečových segmentů, gramatika, ...).
- Uživatelský profil – informace o uživatelích (charakteristiky hlasu, vyjadřování, používané fráze, ...).
- Doménové znalosti – informace odvoditelné z oblasti dialogového systému (gramatika, ...).
- Kontext dialogu – informace o aktuálním stavu dialogu (krok dialogu, uživatelský vstup, chybovost uživatele, ...).

Historie zpracování a napodobování řeči

Dialogové
systémy

Luděk Bártek

O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika

- před 3 milióny let – Australopitekus – schopnost artikulované řeči
- starověk – budování mluvících soch bůžků
- 1779 – Kratzenstein – systém rezonátorů pro napodobení samohlásek a, e, i, o, u.



Historie zpracování a napodobování řeči

Dialogové
systémy

Luděk Bártek

O předmětu

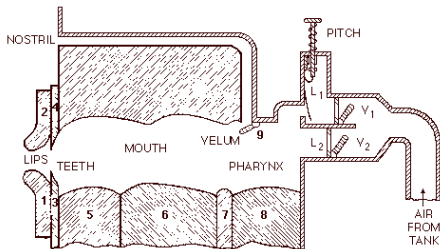
Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika

- 1937 – R. R. Riesz – mechanický mluvící stroj napodobující lidské řečové ústrojí



- 1939 – H. Dudley – VODER (elektromechanický řečový syntetizér), VOCODER (elektrické zařízení kódování a přenos řeči)
- 50. léta 20. století – syntéza ve frekvenční oblasti, později syntéza v časové oblasti

Historie zpracování a napodobování řeči

Dokončení

Dialogové
systémy

Luděk Bártek

O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika

- cca 1970 – počítače
- 1966 – J. Weizenbaum – Eliza (Communications of the ACM, leden).

Základy moderní analýzy řeči

Dialogové
systémy

Luděk Bártek

O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika

- 19. století
 - J.B. Fourier – Fourierova věta – využívá se při spektrální analýze zvuku
 - H. Helmholtz – zabýval se fyziologií vnímání hudby, Helmholtzův rezonátor
 - J. R. Ewald – fyziologie sluchu
- 1924 – spektrální analýza řeči na bázi formantové analýzy samohlásek
- 1939 – vokodér – zařízení pro kompresi řeči pro účely přenosu hlasu rádiem a transkontinentálním kabelem
- 1946 - 1947 – zařízení pro grafický záznam řeči
- 2. polovina 20. století – intenzivní vývoj jak teorie, tak počítačových aplikací.

Základní řečové technologie

Dialogové
systémy

Luděk Bártek

O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika

- syntéza řeči
- rozpoznávání řeči
- související oblasti – zpracování jak v časové tak frekvenční oblasti
- rozpoznávání řečnicka
- detekce emocí
- word spotting
- ...

- Zvuk
 - kmitavý pohyb molekul prostředí (vzduchu)
 - vyvoláván pružným odporem prostředí
- Kmit hmotného bodu
 - pohyb bodu z rovnovážné polohy do místa s maximální výchylkou (amplitudou), odtud do protilehlého místa s maximální výchylkou zpět do rovnovážného bodu.

Kmity

Dialogové
systémy

Luděk Bártek

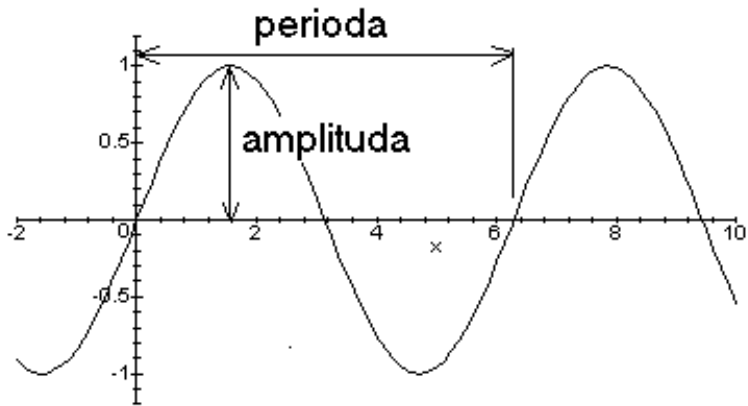
O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika



- Amplituda – maximální výchylka kmitavého pohybu.
- Perioda (T)
 - doba jednoho opakování periodického děje.
 - jednotka – 1 s (sekunda).
- Frekvence (f)
 - počet opakování periodického děje za jednotku času.
 - platí $f = \frac{1}{T}$
 - jednotka 1 Hz (Hertz).

- Síla působící na kmitající bod:
 - $F = -ks$, k – tuhost pružiny, s – aktuální výchylka pružiny
 - $F = ma \Rightarrow ma = -ks$, m – hmotnost tělesa, a – zrychlení
 - $a + \omega^2 s = 0$ ($\omega^2 = \frac{k}{m}$, ω – úhlová rychlost kmitavého pohybu: $\omega = \frac{2\pi}{T}$)
- fáze kmitavého pohybu: $\psi = \omega t$
- okamžitá výchylka: $y = y_m \sin \omega t = y_m \sin \psi$
- okamžitá rychlost: $v = \omega y_m \sin \omega t = y_m \sin \psi$
- okamžité zrychlení: $a = -\omega y_m \sin \omega t = y_m \sin \psi$

Harmonické versus tlumené versus vynucené kmitání

Dialogové
systémy

Luděk Bártek

O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika

- Harmonické kmitání
 - na těleso nepůsobí žádná vnější síla
 - v praxi se s ním téměř nesetkáme (odpor vzduchu, ...).
- Tlumené kmitání
 - proti pohybu působí odpor prostředí
 - amplituda s časem (vzdáleností od zdroje) klesá
- Vynucené kmitání, rezonance
 - na hmotný bod působí navíc periodicky proměnné síla
 $G = \sin \alpha t$
 - $F = ma = -ky + \sin \alpha t \Rightarrow a + \omega^2 y = \sin \alpha t$
 - partikulární řešení: $\frac{\sin \alpha t}{\omega^2 - \alpha^2}$

- Zvuk – mechanické vlnění pružného prostředí (vzduch, voda, kov, ...)
- Akustika – věda studující zvuk (z řeckého akustikos – vztahující se k slyšení):
 - fyzikální – zvuk jako fyzikální vlnění
 - fyziologická akustika – vzniká a vnímání zvuku člověkem
 - hudební – zvuky z pohledu hudby
 - molekulární – vztah akustických vlastností a molekulární struktury.
- Rozdělení zvuku:
 - infrazvuk – frekvence < 16 Hz
 - slyšitelný zvuk – 16 Hz – 16kHz
 - ultrazvuk – > 16 kHz
 - hyperzvuk – až 10^8 Hz – využíván např. molekulární akustikou.

Jednoduchý vs. složený tón

Dialogové
systémy

Luděk Bártek

O předmětu

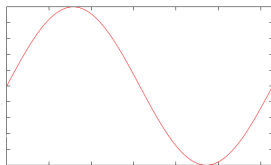
Úvod do
dialogových
systémů

Historie
zpracování
zvuku

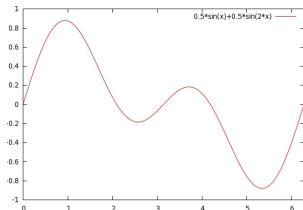
Fyzikální
akustika

Fyziologická
akustika

- Základní tón – průběh intenzity v čase lze popsat jednoduchou sinusoidou.



- Složený tón – lineární kombinace jednoduchých tónů.



Akustické spektrum zvuku

Dialogové
systémy

Luděk Bártek

O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika

- Akustické spektrum – množina základních tónů, ze kterých je zvuk složen.
- Získání spektra – Fourierova transformace:
 - $F(x)$ musí splňovat Dirichletovy podmínky
 - periodická funkce s periodou T
 - je na daném intervalu po částech spojitá (nejvýše konečný počet bodů nespojitosti 1. druhu)
 - má nejvýše konečný počet extrémů na daném intervalu
 - definována v krajních bodech daného intervalu:

Akustické spektrum

Výpočet hodnot

Dialogové
systémy

Luděk Bártek

O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika

- Využívá se rozkladu pomocí Fourierovy řady:

$$F(x) = \frac{a_0}{2} + \sum_{i=1}^{\infty} a_i \cos(i\omega x) + b_i \sin(i\omega x)$$

- $\omega = \frac{2\pi}{T}$
- aproximace $F(x)$ je nejlepší při použití hodnot koeficientů a a b :

$$a_k = \frac{2}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} F(x) \cos(kx) dx$$

$$b_k = \frac{2}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} F(x) \sin(kx) dx$$

- Hodnoty spektrálních koeficientů

$$s_k = \sqrt{a_k^2 + b_k^2}$$

Akustické spektrum zvuku

pokračování

Dialogové
systémy

Luděk Bártek

O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika

- Problém – zvuk je periodický pouze na určitých intervalech.
 - analýza na krátkém intervalu, kde se předpokládá, že je periodický.
- Z hlediska fyziologické akustiky – spektrum odpovídá rezonanci odpovídajících vláček Cortiho ústrojí, resp. odpovídající reakci neuronů.

■ Akustický tlak

- Odpovídá síle působící na element plochy v prostředí akustického vlnění.
- Pro sinusovou vlnu platí:

$$p = p_0 \sin(\omega t)$$

- p_0 – maximální akustický tlak v průběhu periody
- ω – úhlová rychlost
- t – čas.

Akustická intenzita a akustický tlak

Dialogové
systémy

Luděk Bártek

O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika

■ Akustická intenzita

- Vyjadřuje množství akustické energie, které projde jednotkovou plochou za jednotku času.
- Je přímo úměrná druhé mocnině akustickému tlaku.
- Rozsah intenzity zvuku - dán rozsahem minimální (I_0) a maximální (I_1) akustické intenzity, kdy jsme schopni vnímat tón o frekvenci 1 kHz.
- Práh citlivosti – $p_0 = 2 \cdot 10^{-2} \text{ Nm}^{-2}$.
- Práh bolestivosti – $p_1 = 10^2 \text{ Nm}^{-2}$.
- Rozsah – $2,5 \cdot 10^{13} \text{ Nm}^{-2}$.

- Weber-Fechnerův psychofyzikální zákon
 - Člověkem subjektivně vnímaná hlasitost roste při geometrickém nárůstu intenzity přibližně lineárně.
 - Pro stanovení hladiny intenzity zvuku (L) volíme

$$L = 10 \cdot \log \frac{I}{I_0}$$

- jednotka – 1 bel (originál bell) [B]
- Prakticky se využívá odvozená jednotka decibel [dB] ($10^{-1}B$).

Orientační hodnoty akustické intenzity

Dialogové
systémy

Luděk Bártek

O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika

- šepot – 10 - 20 dB
- tlumený hovor – 35 - 45 dB
- symfonický orchestr – 70 - 90 dB
- rocková hudba – 110 - 130 dB.

Základy fyziologické akustiky

Dialogové
systémy

Luděk Bártek

O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika

- Fyziologická akustika se zabývá:
 - mechanismem vytváření řeči
 - mechanismem vnímání řeči.
- Využívá Helmholtzovu rezonanční teorii.

Helmholtzův rezonátor

Dialogové
systémy

Luděk Bártek

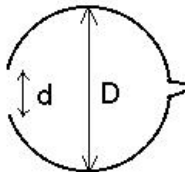
O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika



■ Princip činnosti:

- Přivedením vzduchu do rezonátoru v něm vznikne přetlak.
- Ten vytlačuje přebytečný vzduch ven a následně vzniká podtlak, který způsobí nasávání vzduchu z okolí.
- Takto vzniká periodický děj:

$$f = \frac{75,3}{D} \sqrt{\frac{d}{D}} [\text{Hz}]$$

Mechanismus vytváření řeči

Dialogové
systémy

Luděk Bártek

O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika

- Řeč vzniká pomocí *hlasového ústrojí* (umístěno v *hrtanu*).
- Hlasivky vytváří úzkou hlasovou štěrbinu a jsou rozechvívány procházejícím vzduchem.
- Frekvence jejich kmitání určuje *základní hlasivkový tón* – F_0 .
- Zvuk, který vzniká v hrtanu pomocí hlasivek (samohlásky, znělé souhlásky) je modifikován v *rezonančních dutinách*:
 - hrtanové
 - ústní
 - nosohltanové.
- Rezonanční dutiny fungují na stejném principu jako Helmholtzův rezonátor).

Hlasivky a schéma lidského hlasového ústrojí

Dialogové
systémy

Luděk Bártek

O předmětu

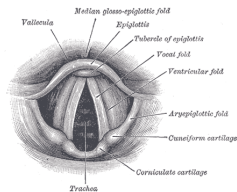
Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika

■ Hlasivky



■ Jejich umístění



- Zvuk vnímáme sluchovým orgánem.
- Sluchový orgán:
 - vnější ucho – zachycuje, soustřeďuje a přivádí zvukové vlny ke střednímu uchu
 - střední ucho
 - mechanickou cestou přenáší zvukovou energii mezi vnějším a vnitřním uchem
 - obsahuje mechanismy k vyrovnání rozdílů tlaku mezi vnějším prostředím a sluchovým orgánem
 - vnitřní ucho – převádí zvukovou energii na vzruchy, které jsou vedeny dále do mozku.

Schéma sluchového orgánu

Dialogové
systémy

Luděk Bártek

O předmětu

Úvod do
dialogových
systémů

Historie
zpracování
zvuku

Fyzikální
akustika

Fyziologická
akustika



Obrázek: Schéma sluchového orgánu

■ Obsahuje:

- Ušní boltec – soustředí zvukové vlny do zvukovodu.
- Zvukovod – vede zachycenou zvukovou energii (vlny) k bubínku.
- Bubínek:
 - Tenká blána na konci zvukovodu – síla cca 0.1 mm.
 - Zesílí a přeneše zvukovou energii na kůstky středního ucha.

- **Obsahuje:**
 - **Kůstky středního ucha:**
 - kladívko – přiléhá k bubínku
 - kovadlinka
 - třmínek – přiléhá k oválnému okénku, kterým se zvuková energie předává do vnitřního ucha.
 - **Oválné okénko** – tvoří přístup k vnitřnímu uchu.
 - **Eustachova trubice:**
 - Vede ze středního ucha do nosohltanu.
 - Slouží k vyrovnání rozdílu tlaku mezi vnějším prostředím a středním uchem, aby nedošlo poškození sluchu.

- Hlemýžď' (Cochlea):
 - Je naplněn vodnatým roztokem.
 - Ústrojí ve tvaru ulity hlemýždě, které obsahuje Cortiho ústrojí.
 - Cortiho ústrojí obsahuje zhruba 20000 vláček s délkami $40 \mu\text{m} - 0,5 \text{ mm}$.
 - Vláčekna jsou jsou napojena na nervová zakončení, která vedou vzruchy do příslušného centra v mozku.
- Rovnovážný orgán.