

AKUSTICKÉ LISTY

České akustické společnosti

ročník 8, číslo 2

červen 2002

Obsah

Prof. Felix Kolmer osmdesátiletý	2
Certifikace v akustice <i>Jindřich Schwarz</i>	2
64. akustický seminář <i>Vlastimil Strnad</i>	2
Úvod do bayesovské klasifikace dat <i>Roman Čmejla a Pavel Sovka</i>	3
Clinical interpretation of the acoustic pressure measurement in the human ear canal <i>Jan Mejzlík, Zdeněk Škvor, Libor Husník, František Rund</i>	11

Prof. Felix Kolmer osmdesátiletý



V květnu se dožívá 80 let prof. Ing. Felix Kolmer, DrSc., stále aktivní čestný člen České akustické společnosti. Patří k těm, kteří se svou dlouholetou odbornou i organizační prací nejvíce zasloužili o rozvoj oboru akustiky v Čechách. Před válkou žil prof. Kolmer v Praze a v Rakousku, v době okupace byl zatčen gestapem a 4 roky vězněn v koncentračních táborech Terezín, Osvětim a Friedland. Je jedním z 3100 vězňů z Československé republiky, který přežil vyhlazovací tábory. Jeho osudy za 2. světové války natočila Spielbergova nadace v Los Angeles a Univerzita v Oldenburgu.

V roce 1949 nastoupil do tehdejšího Výzkumného ústavu zvukové, obrazové a reprodukční techniky v Praze, v němž pracoval až do důchodu. Začal zde jako pracovník elektroakustického oddělení a posléze byl dlouhá léta ředitelem. Vždy věnoval péči podmínkám rozvoje akustiky a to jak ve VÚZORTu, tak i mimo něj. V případě potřeby neváhal použít své autority, aby pomohl při řešení problémů. Přestože v roce 1990 odešel do důchodu, jeho zájem o akustiku neskončil a stále sleduje dění v oboru.

Je autorem nebo spoluautorem 200 publikací, z nichž 24 bylo uveřejněno v zahraničí. Bohatá byla rovněž jeho pedagogická činnost v oboru akustiky. Působil na Fakultě elektrotechnické ČVUT a na FAMU, dosud je členem Státní komise pro závěrečné zkoušky v oborech audiotechnika a zvuková tvorba. Hodně svého času věnoval prof. Kolmer činnosti ve společenských organizacích v oboru akustiky. V roce 1959 pomáhal zakládat prof. Slavíkovi Akustickou komisi ČSAV a řadu let byl jejím předsedou. Byl předsedou Technické normalizační komise pro akustiku od jejího založení v roce 1974 až do odchodu do důchodu a nadále je jejím čestným členem. Dlouhá léta pravidelně reprezentoval Československo na zasedáních technické komise 43 „Akustika“ Mezinárodní organizace pro normalizaci, kde rovněž zastupoval IUPAP (International Union of Pure and Applied Physics). Byl jedním ze zakladatelů Federation of Acoustical Societies of Europe (FASE) a 12 let byl jejím generálním tajemníkem. V International Commission of Acoustics (ICA) byl ředitelem Informačního a koordinačního centra. Je Fellow of Acoustical Society of America, čestným členem několika zahraničních akustických společností a nositelem tuzemských i zahraničních medailí oceňujících jeho dlouholetou práci v oboru. V současnosti se prof. Kolmer věnuje hlavně aktivitám souvisejícím s česko-německými vztahy a odškodněním obětí. Je členem předsednictva Terezínské iniciativy, členem výboru Českých osvětimských vězňů a dalších institucí s podobným zaměřením. Ve službách Ministerstva zahraničí ČR se zúčastňuje v USA a v Německu jednání o odškodnění obětí totálního nasazení a česko-německých vztazích.

Blahopřejeme a přejeme hodně zdraví a energie do dalších let!

Certifikace v akustice

Ve středu 29. května 2002 se sešel přípravný výbor Správní rady Certifikačního orgánu ČsAS, aby pod vedením předsedkyně RNDr. Anny Ryndové, Ph. D. projednal stav příprav certifikace pracovníků v akustických oborech. Správní rada konstatovala, že systém certifikace je v zásadě připraven. Je stanovena organizační struktura certifikačního orgánu, je připravena koncepce Příručky jakosti, strategie přípravy pracovníků i ověřování jejich odborné způsobilosti. Přípravný výbor Správní rady dále připravuje personální zabezpečení certifikace. Přípravný výbor proto hodnotí stav příprav certifikace jako dostatečný pro zahájení jednání o akreditaci certifikačního orgánu Národním akreditačním orgánem, Českým institutem pro akreditaci, o. p. s. Tomu však zatím brání nedostatečné finanční zabezpečení startu certifikace. Správní rada proto požádala Radu ČsAS, aby se obrátila dopisem na ty firmy, o nichž předpokládá, že mají zájem o certifikaci svých odborníků – akustiků, s žádostí o pomoc při finančním zabezpečení startu certifikace.

Jindřich Schwarz

64. akustický seminář

Ve dnech 14. až 16. května 2002 proběhl ve Vysokém nad Jizerou 64. akustický seminář. Týkal se hluku a vibrací se zvláštním zaměřením na stroje, zařízení, dopravní techniku a další příbuznou problematiku. Semináře se zúčastnilo celkem 66 osob, členů i nečlenů ČsAS z ČR, SRN a Rakouska. Předneseno bylo 19 odborných příspěvků z výše uvedených oblastí včetně legislativy. Zástupci firem Brüel & Kjær, LB electronics, NORSONIC, PHOENIX a SYNERMA prezentovali své výrobky a měřicí techniku. Součástí semináře byla i společná schůze odborných skupin ČsAS.

Na základě dostatečné nabídky referátů pro otištění bude dle původního oznámení vydán dodatečně sborník.

Seminář byl doprovázen vysokým zájmem účastníků o diskutovaná témata, čehož dokladem byla i hojná účast na závěrečných přednáškách.

Vlastimil Strnad

Úvod do bayesovské klasifikace dat

Roman Čmejla a Pavel Sovka

ČVUT-FEL, Technická 2, 16627 Praha 6

email: [cmejla, sovka]@feld.cvut.cz

The paper is focused on the introduction into Bayesian classification techniques. Necessary terms are introduced, and basic characteristics of Bayesian based classification are discussed. Especially, the influence of apriori information is illustrated. Also the use of Bayesian evidence for the proper model choice is discussed. As illustrative examples the classification of speech are used, e.g. the simple classification of occlusion and vocals using discrimination function. Gained results are compared with results given by classical approach based on Euclidean metric. Examples of the application in biological signal processing are also given.

1. Úvod

Článek uvádí do problematiky bayesovské klasifikace dat pomocí diskriminační analýzy. Jedná se o základní a často používaný postup klasifikačních procedur. Podrobně je popsán význam a použití jednotlivých členů Bayesova vztahu. Pozornost je věnována vlivu apriorní informace na výsledek klasifikace. Rovněž je diskutován význam bayesovské evidence pro volbu modelu a význam pro segmentaci řeči a zpracování biologických dat. Jako příklad použití bayesovských technik je vybrána klasifikace řečových dat. Na příkladech jednoduché klasifikace exploziv a klasifikace samohlásek jsou ukázány možnosti použití diskriminační funkce a její vlastnosti. Získané výsledky jsou porovnány s klasickým přístupem založeným na použití vzdálenosti v euklidovské metrice. Uvedené příklady byly vybrány pro svoji jednoduchost, názornost a snadnou realizovatelnost. Lze na nich dobře ilustrovat vlastnosti jednoduchých klasifikátorů.

2. Vymezení použitého přístupu klasifikace řečových dat

Pro klasifikaci dat, která je častou úlohou v mnoha aplikacích, existuje velmi mnoho metod odvozených pomocí deterministického nebo pravděpodobnostního přístupu. V dalším textu se budeme věnovat především klasifikaci řečových dat. Tato problematika je důkladně popsána např. v [13]. Cílem tohoto příspěvku je na vybraném příkladě klasifikace fonetických jednotek řeči uvést základní principy a vlastnosti přístupu založeného na Bayesově vzorci.

Z akusticko-fonetického pohledu lze pro klasifikaci řeči do fonetických tříd použít různých akustických parametrů, které jsou odhadovány pro krátké segmenty řeči. K často používaným patří např. průchody signálu nulou, energie, základní frekvence řeči, rezonanční frekvence hlasového traktu (formanty) nebo kepstrální koeficienty¹. Lze samozřejmě použít rovněž kombinace těchto parametrů, tedy

např. formanty a energii. V tomto příspěvku, vzhledem k názornosti, použijeme ke klasifikaci samohlásek (A, E, I, ...) pouze první dva formanty a pro klasifikaci frikativ (R, V, F, ...) tři kepstrální koeficienty. Nicméně pro teoretický výklad budeme pracovat s M rozměrným vektorem parametrů stejného typu nebo jejich kombinací.

Máme-li fonetické kategorie (třídy) popsané množinou M parametrů, pak nejjednodušším rozhodnutím o zařazení naměřených dat do těchto tříd je rozhodnutí na základě euklidovské vzdálenosti

$$v = \sqrt{\sum_{m=1}^M (p[m] - q[m])^2}, \quad (1)$$

kde $p[m]$ představuje M parametrů referenčního vzoru a $q[m]$ je M parametrů testovaného úseku signálu (vzorku). Je-li vzdálenost v menší než vhodně určený práh, přiřadíme testovaný vzorek k danému referenčnímu vzoru, přesněji do třídy určené tímto referenčním vzorem. Vzdálenost v totiž nebývá nulová, a proto není testovaný vzorek roven referenčnímu vzoru, ale je mu blízký. Říkáme proto, že testovaný vzorek patří do třídy, jejíž vlastnosti jsou určeny vzorem a velikostí použitého práhu. Tento přístup náleží k deterministickým metodám a neumožňuje efektivně využít informaci obsaženou v signále. Spolehlivější a robustnější způsob klasifikace získáme využitím pravděpodobnostního počtu.

Pro další text je vhodné vymežit pojem model. Budeme jej používat ve dvou významech. Jednak jím budeme označovat typ parametrizace použité pro popis dat, jednak jej budeme používat pro označení referenčního vzoru získaného z dat. Oba významy budou zřejmé z kontextu. Pro řeč často používáme parametry získané spektrální analýzou (např. formantové kmitočty), kepstrální analýzou nebo lineárním predikčním (LP) kódováním, např. [13], [14], [15], [7]. Pro analýzu biologických signálů jsou nejčastěji používané spektrální a LP modely.

¹Ty lze získat např. zpětnou Fourierovou transformací logaritmického spektra segmentu signálu. Tyto koeficienty proto nesou informaci o tvaru logaritmického spektra.

3. Bayesův vzorec

Z teorie pravděpodobnosti vyplývá pravidlo pro násobení podmíněných a nepodmíněných pravděpodobností jevů A a B , např. [1], [10], [11]

$$P(B) \cdot P(A|B) = P(A) \cdot P(B|A), \quad (2)$$

kde podmíněnou pravděpodobností jevu $P(A|B)$ rozumíme pravděpodobnost, že nastane jev A za podmínky, že nastal jev B . Naopak $P(B|A)$ představuje pravděpodobnost, že jev B nastane za podmínky existence jevu A .

Nyní vzorec (2) použijme pro popis situace, kdy existuje větší počet jevů. Místo označení A zavedme označení h_i , které představuje hypotézu, že testovaný vzorek náleží do i -té třídy dané referenčním vzorem. Pro popis existence jevu B , tedy pro situaci, kdy získáme vektor parametrů popisující naměřená data, zavedme potom označení \mathbf{d} . Místo velkých písmen označující pravděpodobnosti (tedy čísla) použijeme malá písmena označující hustoty pravděpodobnosti, tedy funkce vektoru parametrů \mathbf{d} . Označení $p(h_i|\mathbf{d})$ potom představuje situaci, kdy stanovujeme pravděpodobnost pravdivosti hypotézy h_i za podmínky, že parametry signálu nabyly určitých hodnot. Výše uvedené pravidlo (2) můžeme přepsat do tvaru, který nazýváme *Bayesovým vzorcem* (anglický matematik Thomas Bayes 1702 - 1761)

$$p(h_i|\mathbf{d}) = \frac{p(\mathbf{d}|h_i) \cdot P(h_i)}{p(\mathbf{d})} = \frac{p(\mathbf{d}|h_i) \cdot P(h_i)}{\sum_{i=1}^S p(\mathbf{d}|h_i) \cdot P(h_i)}, \quad (3)$$

kde S je počet tříd, do kterých naměřená data při klasifikaci třídíme. Počet tříd je samozřejmě roven počtu hypotéz, které testujeme, proto je použit symbol h_i . V uvedeném vztahu se vyskytuje jednak pravděpodobnost $P(h_i)$, která představuje číslo, jednak i hustoty pravděpodobností $p(h_i|\mathbf{d})$, $p(\mathbf{d}|h_i)$ a $p(\mathbf{d})$, které jsou funkcemi parametrů \mathbf{d} dat². Levá strana uvedeného vztahu $p(h_i|\mathbf{d})$ vyjadřuje pravděpodobnost hypotézy, že naměřená data s parametry \mathbf{d} patří do i -té třídy. Význam vztahu spočívá ve skutečnosti, že zatímco hustotu pravděpodobnosti $p(h_i|\mathbf{d})$ neznáme, hustoty pravděpodobnosti na pravé straně vztahu (3) určit umíme, máme-li k dispozici data. Abychom si tuto skutečnost mohli podrobněji vysvětlit, popíšeme podrobněji jednotlivé členy Bayesova vzorce.

- Pravděpodobnost $P(h_i)$ se nazývá *apriorní pravděpodobnost hypotéz* a představuje míru pravdivosti hypotézy h_i . V aplikacích Bayesova vzorce se hypotézy h_i navzájem vylučují a právě jedna je správná. Jsou to pravděpodobnosti hypotéz *před pozorováním* signálu a někdy jsou známy. Pokud ne, tak jedním z možných je zavedení předpokladu, že hypotézy jsou

²V případě jednorozměrného normálního rozdělení [7] těchto hustot pravděpodobností je lze popsat dvěma parametry, a to střední hodnotou μ a směrodatnou odchylkou σ . Označujeme je potom jako $N(\mu, \sigma)$.

skutečně apriori stejně možné, a tedy jim přiřadíme stejnou pravděpodobnost. Takový přístup bývá označován Bayesovým postulátem. Někdy lze použít i subjektivního odhadu apriorní pravděpodobnosti. Z existence tohoto členu Bayesova vzorce vyplývá zásadní rozdíl mezi bayesovskými technikami a metodou maximální věrohodnosti označované jako ML (maximum likelihood). V některých případech může metoda maximální věrohodnosti selhat, neboť nepoužívá informaci danou právě touto apriorní pravděpodobností.

- Podmíněná hustota pravděpodobnosti $p(\mathbf{d}|h_i)$ je nazývána *pravděpodobnostní mírou* výsledku testu na datech \mathbf{d} za předpokladu platnosti hypotézy h_i . Pokud zvažujeme různé hypotézy pro stejný soubor dat, pak tento výraz vyjadřuje *věrohodnostní funkci* (často používanou i v nebayesovských úlohách, např. již zmíněné metodě maximální věrohodnosti). Zvažujeme-li pro jednu hypotézu různé soubory dat, pak tato pravděpodobnost představuje *vzorkovací distribuci*. V úlohách uvedených v tomto článku jsou testovány různé hypotézy pro stejný soubor dat, a proto budeme pro pravděpodobnostní míru rovněž používat termín věrohodnostní funkce, nebo zkráceně věrohodnost. Hustotu pravděpodobnosti $p(\mathbf{d}|h_i)$ umíme získat statistickou analýzou dat, která máme k dispozici. V praxi tuto hustotu můžeme aproximovat histogramem vybraného parametru dat. Příklad bude uvedený v dalším textu.
 - Hustota pravděpodobnosti $p(\mathbf{d})$ je nazývána *úplnou pravděpodobností* (někdy též *evidencí* [8]). Tento člen Bayesova vzorce je vždy začleněn do normovací konstanty a neovlivňuje závěry učiněné v rámci vybraného modelu dat. Protože však ovlivňuje relativní pravděpodobnost různých modelů, bývá předmětem našeho zájmu při jejich výběru. Tento případ je typický např. pro segmentaci řeči [4], [3], [5] nebo pro hledání optimálního modelu záznamu mozkové aktivity (EEG signálu). Při výběru vhodného modelu lze velmi přesně určit náhlé změny v EEG signálu snímaného na povrchu hlavy. Přesnost a spolehlivost detekce je srovnatelná s výsledky získanými pomocí elektrod vnořených do mozkové tkáně. Proto je tento problém v současné době intenzivně zkoumán. Výběr modelu se rovněž uplatňuje při potlačování šumů při mapování mozkové aktivity na povrch hlavy a lokalizaci aktivních částí mozku [9].
- Zde se opět bayesovské techniky odlišují od klasických statistických metod, které problém výběru vhodného modelu z dané množiny modelů neřeší [6].
- Podmíněná hustota pravděpodobnosti $p(h|\mathbf{d})$ je *aposteriorní pravděpodobnost* určená daty a apriorní informací. Vyplývá z přímého použití Bayesova vzorce a ukazuje nám, jak pravděpodobná je hypotéza h_i *po pozorování* signálu \mathbf{d} .

4. Diskriminační funkce a klasifikace

Nyní je před námi úloha, kdy chceme pro pozorovaná data potvrdit či zamítnout hypotézu o jejich příslušnosti do některé třídy. Základním problémem je nalezení rozhodovacího pravidla, pomocí kterého na základě porovnání daného vektoru parametrů pozorovaných dat s vektory parametrů referenčních vzorů (s modely) provedeme přiřazení dat do tříd (tedy ověření jednotlivých hypotéz).

Jako příklad uvažujme jednorozměrný problém, ve kterém jsou předmětem klasifikace data popsána konkrétní hodnotou d_0 jednoho parametru d , která je třeba třídit do dvou tříd (dvě hypotézy).

Naměřená data přiřadíme do třídy 1 (hypotéza h_1), pokud pro hodnoty aposteriorních pravděpodobností platí³

$$p(h_1|d) > p(h_2|d), \quad \text{pro } d = d_0. \quad (4)$$

Nerovnost lze s použitím vztahu 3 přepsat do tvaru

$$\frac{p(d|h_1)P(h_1)}{p(d)} > \frac{p(d|h_2)P(h_2)}{p(d)}. \quad (5)$$

Protože hustota pravděpodobnosti $p(d)$ se vyskytuje na obou stranách nerovnice, můžeme pravidlo pro přiřazení dat do třídy 1 (hypotéza h_1) přepsat do tvaru

$$p(d|h_1)P(h_1) > p(d|h_2)P(h_2). \quad (6)$$

Často bývají apriorní pravděpodobnosti stejné $P(h_1) = P(h_2)$. V těchto případech je v nerovnosti můžeme vynechat, čímž získáme výsledný vztah pro klasifikaci pomocí pravděpodobnostní míry (věrohodnostní funkce, věrohodnosti)

$$p(d|h_1) > p(d|h_2). \quad (7)$$

Vybíráme tedy tu hypotézu, jejíž pravděpodobnostní míra (věrohodnost) je pro danou konkrétní hodnotu d_0 parametru d větší. Bod (hodnota parametru d), ve kterém se obě hodnoty obou pravděpodobnostních měr (věrohodností) rovnají $p(d|h_1) = p(d|h_2)$, se nazývá *rozhodovací hranicí*. Pro jednorozměrné případy se jedná o jediný bod (viz obr. 1a), avšak pro více dimenzí se jedná o křivku či povrch v prostoru. Na obr. 1a jsou pravděpodobnostní hustoty $p(d|h_1)$ a $p(d|h_2)$ vyneseny pro jediný parametr d , který v tomto případě představuje délku trvání exploziv v milisekundách. Proto je průsečíkem křivek jediný bod, jemuž odpovídá hodnota parametru $d = 116$ ms. Vlevo od této hodnoty platí $p(d|h_1) > p(d|h_2)$, vpravo od této hodnoty naopak platí $p(d|h_1) < p(d|h_2)$.

Pro rozhodovací pravidla se často zavádí *diskriminační funkce* $\tilde{g}_i(d)$ pro hypotézu h_i

$$\tilde{g}_i(d) = p(d|h_i)P(h_i) \quad (8)$$

a rozhodovací pravidlo přechází na výběr hypotézy h_1 při splnění podmínky $\tilde{g}_1(d) > \tilde{g}_2(d)$ (pro dané $d = d_0$) a h_2 při nesplnění této podmínky.

Často se pro vyjádření diskriminační funkce používá logaritmy

$$g_i(d) = \ln(p(d|h_i)) + \ln(P(h_i)), \quad (9)$$

čímž násobení pravděpodobností převedeme na součet jejich logaritmy. To má význam jednak pro numerickou stabilitu algoritmů, jednak pro realizaci velkých objemů výpočtů.

5. Příklad jednoduché klasifikace

Příklad použití Bayesova vztahu si ukážeme na jednoduchém příkladě klasifikace. Předpokládejme, že máme k dispozici dokonalý identifikátor všech okluzí⁴ v plynulé řeči. S použitím Bayesova vzorce (3) se pokusíme rozdělit všechny okluze do dvou základních kategorií - znělá/neznělá okluze. Pro klasifikaci okluzí si vybereme jediný parametr: časové trvání (délku) okluze.

Parametr d ve vztazích (4) až (9) reprezentuje časové trvání jednotlivých okluzí. Hypotéza h_1 odpovídá fonetická kategorie znělá okluze a hypotéza h_2 kategorie neznělá okluze. Apriorní pravděpodobnosti $P(h_i)$ určíme pomocí údajů o relativní frekvenci českých fonémů, které jsou uvedené v pramenu [13]. Apriorní pravděpodobnost pro znělé okluze {B, D, Ď, G} má hodnotu $P(h_1) = 0,33$, pro neznělé okluze {P, T, Ě, K} má hodnotu dvakrát větší $P(h_2) = 0,67$.

Abychom mohli co nejpřesněji odhadnout podmíněnou pravděpodobnost (věrohodnost) $p(d|h)$, potřebovali bychom co největší možné množství realizací obou typů okluzí s různou dobou trvání. Ve skutečnosti však pracujeme pouze s malým počtem (řádově desítky až stovky) realizací, které dávají nepřesné odhady pravděpodobnostního rozdělení. V práci [2] jsou na str. 53 a 54 uvedeny histogramy trvání znělých a neznělých okluzí. Ačkoliv dávají pouze hrubou aproximaci (správné rozdělení by se mělo blížit hladké křivce), můžeme z jejich tvaru učinit předpoklady o rozdělení pravděpodobnosti. Podobně jako většina charakteristik používaných při analýze řeči (formantové frekvence, průchody nulou, amplitudy na výstupu banky filtrů, apod.) splňují i délky okluzí předpoklad normálního (gaussovského) rozdělení. Věrohodnost přiřazení dané realizace okluze do jedné ze dvou uvedených skupin (hypotéza h_1 představuje přiřazení okluze do skupiny znělých okluzí, h_2 přiřazení k neznělým okluzím) má potom tvar

$$p(d|h) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(d-\mu)^2}{2\sigma^2}}, \quad (10)$$

kde směrodatná odchylka

³Funkce $p(h_1|d)$ a $p(h_2|d)$ v nerovnosti 4 jsou vyčísleny pro konkrétní hodnotu parametru d_0 , představují proto v tomto případě jedno číslo.

⁴Okluze představují pauzu před explozí, se kterou tvoří jeden celek nazývaný explozíva. Neznělá okluze představuje pauzu o délce přibližně 120 ms, znělá okluze periodický signál o frekvenci přibližně 200 Hz o malé amplitudě a délce přibližně 80 ms.

$$\sigma = \sqrt{\frac{1}{N} \sum_{j=0}^N (d_j - \mu)^2} \quad (11)$$

představuje míru „rozprostření“ dat.

Z histogramů [2] získáme $p(d|h_1) = N(88, 42; 21, 28)$ a $p(d|h_2) = N(142, 5; 30, 99)$.

Úplná pravděpodobnost dat (evidence) je dána jmenovatelem vztahu (3)

$$p(d) = p(d|h_1) \cdot P(h_1) + p(d|h_2) \cdot P(h_2) \\ = 0,33 N(88, 42; 21, 28) + 0,67 N(142, 5; 30, 99)$$

Aposteriorní pravděpodobnosti jsou podle (3)

$$p(h_1|d) = \frac{0,33 N(88, 42; 21, 28)}{0,33 N(88, 42; 21, 28) + 0,67 N(142, 5; 30, 99)},$$

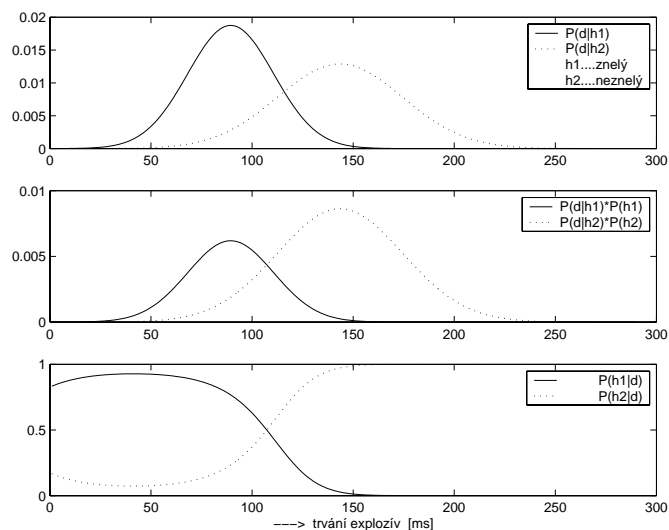
$$p(h_2|d) = \frac{0,67 N(142, 5; 30, 99)}{0,33 N(88, 42; 21, 28) + 0,67 N(142, 5; 30, 99)}.$$

Při klasifikaci postupujeme tak, že pro danou délku délku okluze d pomocí vztahu (3) vypočítáme hodnoty aposteriorních pravděpodobností $p(h_1|d)$ a $p(h_2|d)$. Je-li $p(h_1|d) > p(h_2|d)$, pak zařadíme testovanou okluzi do skupiny znělých okluzí (platí hypotéza h_1). Naopak, je-li $p(h_1|d) < p(h_2|d)$, pak zařadíme testovanou okluzi mezi neznělé okluze (platí hypotéza h_2). Klasifikujeme tedy na základě maxima aposteriorní pravděpodobnosti. Tento postup je v literatuře označován jako MAP (maximum aposteriori) kritérium.

Odlišnost popsaného postupu od postupu klasifikace založené na metodě maximální věrohodnosti (využívající věrohodnost místo aposteriorní pravděpodobnosti) ilustrují průběhy na obrázku 1. Tento obrázek tedy ilustruje vzájemný vztah aposteriorní pravděpodobnosti $p(h_i|d)$ a věrohodnostní funkce $p(d|h_i)$.

Na obr. 1a jsou zakresleny věrohodnosti $p(d|h_1)$ a $p(d|h_2)$ používané v metodě maximální věrohodnosti. Jejich průsečík představuje rozhodovací mez, která je v našem případě 116 ms. V prostředním obrázku jsou zobrazeny součiny věrohodností násobených apriorními pravděpodobnostmi $P(h_i)$. V tomto případě se rozhodovací mez snížila na 108 ms. Průběhy na obr. 1c představují hustoty aposteriorních pravděpodobností $p(h_i|d)$. Vidíme, že rozhodovací mez zůstává stejná jako v předcházejícím případě, neboť evidence $p(d)$ ze vztahu (3) je v obou případech stejná.

Zvolíme-li jako příklad délku okluze 110 ms, pak při rozhodování na základě maximální věrohodnosti (tedy s pomocí věrohodnosti $p(d|h_1)$ a $p(d|h_2)$) ji zařadíme mezi znělé okluze, zatímco s použitím bayesovské metody (tedy s pomocí aposteriorních pravděpodobností $p(h_1|d)$ a $p(h_2|d)$) do neznělých okluzí.



Obrázek 1: a) rozdělení podmíněných pravděpodobností (pravděpodobnostních měr, věrohodností) $p(d|h_i)$ b) součin věrohodností $p(d|h_i)$ a apriorních pravděpodobností $P(h_i)$ c) aposteriorní pravděpodobnosti $p(h_i|d)$

6. Vícerozměrný případ

Postup uvedený v předchozí kapitole nyní zobecníme na případ, kdy jsou data charakterizována vektory s více parametry. Rozhodovací hranicí ve vícerozměrném případě je v tomto případě povrch, který dělí prostor mezi různými typy tříd. Diskriminační funkce ve vícedimenzionálním případě používá vícerozměrného gaussovského rozdělení [7]

$$p(\mathbf{d}|h_i) = (2\pi)^{-\frac{M}{2}} |\mathbf{C}_i|^{-\frac{1}{2}} e^{-\frac{(\mathbf{d}-\mu_i)' \mathbf{C}_i^{-1} (\mathbf{d}-\mu_i)}{2}}, \quad (12)$$

kde $|\mathbf{C}_i|$ je determinant kovarianční matice \mathbf{C}_i a μ_i je vektor středních hodnot M parametrů pro i -tou třídu (hypotézu h_i). Vektor \mathbf{d} obsahuje M parametrů testovaného vzoru.

Zjednodušená diskriminační funkce předpokládající rovnoměrné apriorní rozložení je logaritmem vícerozměrného gaussovského rozdělení

$$g_i(\mathbf{d}) = \ln(\mathbf{d}|h_i) = -\ln |\mathbf{C}_i| - (\mathbf{d}-\mu_i)^T \mathbf{C}_i^{-1} (\mathbf{d}-\mu_i), \quad (13)$$

kde konstantní první člen rovnice (12) je vynechán. Druhý člen funkce $g_i(\mathbf{d})$ je znám jako často používaná Mahalanobisova vzdálenost

$$D_i^2(\mathbf{d}) = (\mathbf{d}-\mu_i)^T \mathbf{C}_i^{-1} (\mathbf{d}-\mu_i). \quad (14)$$

Při ověřování platnosti hypotézy h_i používáme maxima funkcí $p(\mathbf{d}|h_i)$ a $g_i(\mathbf{d})$ a naopak minimum funkce $D_i^2(\mathbf{d})$. V prvních dvou případech hledáme maximální pravděpodobnost, že testovaný vzorek náleží k určité třídě, zatímco

ve třetím případě hledáme minimální vzdálenost testovaného vzorku od určitého třídy reprezentované vhodným modelem.

Postup určení μ_i a \mathbf{C}_i bude vysvětlen na příkladě v následujícím odstavci.

7. Příklady vícerozměrné klasifikace

Celý klasifikační proces zahrnuje dvě fáze – trénovací a testovací. V trénovací fázi musíme vytvořit modely jednotlivých tříd. V případě normálního rozdělení je třeba určit střední hodnoty μ_i a kovarianční matice \mathbf{C}_i pro jednotlivé třídy. Dvojice μ_i a \mathbf{C}_i představuje model i -té třídy, se kterým porovnáváme testovaný vzorek. Podle použitého kritéria hledáme buď maximální pravděpodobnost (12) a (13) nebo minimální vzdálenost (14). Používáme-li v trénovací a testovací fázi různá data, pak hovoříme o otevřených testech. Jsou-li trénovací a testovací data shodná, jedná se o uzavřený test. Výsledky testů se často udávají pomocí konfuzních tabulek či matic. Je zřejmé, že uzavřený test poskytuje lepší výsledky (v ideálním případě 100% úspěšnost) než otevřený test, a proto jej lze použít k odhadu horní hranice úspěšnosti zvolené metody. K této hranici se při praktickém použití klasifikační metody lze blížit zdola, neboť reálné použití nejčastěji odpovídá podmínkám otevřeného testu. Otevřený test umožňuje určit, jak se daná klasifikační metoda vyrovná s neznámými daty, tedy jakou má schopnost zobecnění (generalizace). Z obou uvedených důvodů jsou pro měření úspěšnosti klasifikace častěji používány otevřené testy než testy uzavřené.

7.1. Příklad klasifikace samohlásek

A. Zadání

Uvažujme 44 realizací každé z pěti samohlásek A, E, I, O, U. Každá samohláska je popsána 44 páry formantových frekvencí F1a F2 [Hz], které tvoří data (přesněji vektor parametrů dat) \mathbf{d} .

Pro ilustraci uvedeme příklad formantových kmitočtů pro samohlásku A:

Frekvence prvních formantů jsou dány 44 čísla

$F1_A = [742; 645; 613; 645; 742; 645; 710; 710; 710; 613; 774; 806; 645; 710; 806; 774; 742; 806; 839; 839; 806; 839; 903; 871; 871; 968; 742; 871; 710; 774; 935; 645; 935; 806; 968; 710; 935; 871; 645; 903; 903; 903; 903; 968]$.

Frekvence druhých formantů jsou vyšší

$F2_A = [1290; 1355; 1290; 1290; 1194; 1226; 1290; 1194; 1097; 1161; 1194; 1161; 1129; 1290; 1194; 1419; 1742; 1548; 1548; 1387; 1419; 1516; 1613; 1516; 1581; 1355; 1677; 1226; 1355; 1581; 1323; 1419; 1484; 1516; 1194; 1484; 1194; 1323; 1581; 1581; 1419; 1516; 1516; 1516]$.

To znamená, že pro třídu samohlásek A jsou jednotlivé prvky $d_k[i]$ datových vektorů \mathbf{d} určeny:

$d_1[1] = 742, d_1[2] = 1290, d_2[1] = 645, d_2[2] = 1355, \dots,$
až $d_{44}[1] = 968, d_{44}[2] = 1516$.

Označení $d_k[m]$, $k = 1, 2, \dots, 44$; $m = 1, 2$ představuje k -tý vzorek z celkového počtu $N = 44$, index m představuje m -tý parametr z celkového počtu $M = 2$. Počet tříd je roven 5, tedy $i = 1, 2, \dots, 5$.

B. Trénovací fáze

Vytvoření modelů pro tento příklad se skládá ze dvou kroků.

- Výpočet vektorů μ_i , $i = 1, 2, \dots, 5$ středních hodnot dvou parametrů (formantové kmitočty). Tím získáme pět dvouprvkových vektorů $\mu_i = [\mu_i[1], \mu_i[2]]$, $i = 1, 2, \dots, 5$. Příslušnost vektoru k dané třídě (dané samohlásce) lze symbolicky vyjádřit označením $\mu_A, \mu_E, \mu_I, \mu_O, \mu_U$.
- Výpočet kovariančních matic \mathbf{C}_i , $i = 1, 2, \dots, 5$ pro pět tříd. Pro výklad použijeme místo \mathbf{C}_i označení $\mathbf{C}_A, \mathbf{C}_E, \mathbf{C}_I, \mathbf{C}_O, \mathbf{C}_U$, ze kterého je zřejmé k jaké třídě (samohlásce) matice náleží.

Výpočet středních hodnot $\mu_i[1]$ a $\mu_i[2]$ pro i -tou třídu je proveden pomocí vztahu

$$\mu_i[k] = \frac{1}{N} \sum_{k=1}^N d_k[i], \quad i = 1, 2, \dots, 5, \quad k = 1, 2. \quad (15)$$

Výsledné hodnoty jsou pro všech pět tříd uvedeny v tabulce 1.

Třída	$\mu_i[1]$	$\mu_i[2]$
A	793	1384
E	615	1866
I	396	1992
O	555	1001
U	400	862

Tabulka 1: Střední hodnoty prvních dvou formantů českých samohlásek

Dvourozměrná (máme 2 parametry) kovarianční matice má tvar

$$C = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}. \quad (16)$$

Na hlavní diagonále leží rozptyly jednotlivých parametrů, mimo hlavní diagonálu jsou vzájemné korelační koeficienty obou parametrů.

Korelační koeficienty jsou definovány vztahem

$$\sigma_{ml} = \frac{N \sum_{k=1}^N d_k[m] d_k[l] - \sum_{k=1}^N d_k[m] \sum_{k=1}^N d_k[l]}{N(N-1)}, \quad (17)$$

kde $d_k[m]$ $m = 1, 2$; $l = 1, 2$ je, jak bylo uvedeno, m -tý parametr k -tého vzorku.

Prvky kovariančních matic pro jednotlivé třídy samohlásek jsou v našem případě rovny

$$\mathbf{C}_A = \begin{bmatrix} 11486 & 4558 \\ 4558 & 27288 \end{bmatrix}; \quad \mathbf{C}_E = \begin{bmatrix} 9602 & 4004 \\ 4004 & 25697 \end{bmatrix};$$

$$\mathbf{C}_I = \begin{bmatrix} 6609 & -4563 \\ -4563 & 35109 \end{bmatrix}; \quad \mathbf{C}_O = \begin{bmatrix} 5818 & 796 \\ 796 & 7454 \end{bmatrix};$$

$$\mathbf{C}_U = \begin{bmatrix} 6499 & -989 \\ -989 & 12847 \end{bmatrix}.$$

C. Testovací (klasifikační) fáze

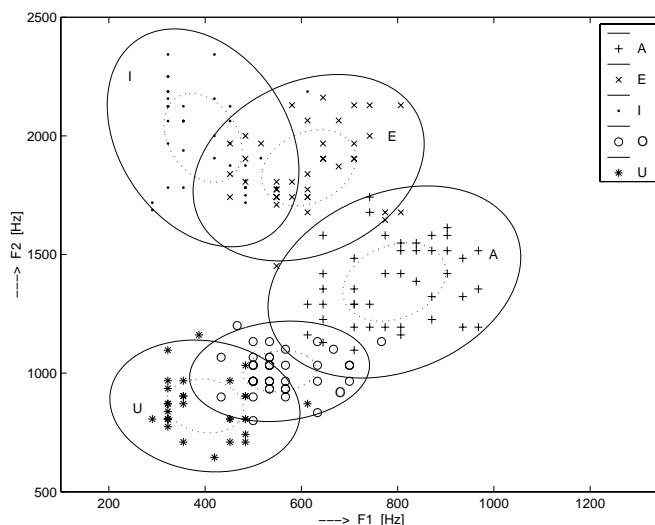
Při klasifikaci testujeme příslušnost vybraného vzorku popsaného parametry vektoru \mathbf{d} dosazováním hodnot vektoru \mathbf{d} do jednoho ze vztahů (12)-(14) pro všechny třídy $i = 1, 2, \dots, 5$. Tím získáme 5 hodnot veličin $p(\mathbf{d}|h_i)$, $g_i(\mathbf{d})$ nebo $D_i^2(\mathbf{d})$. Největší hodnota $p(\mathbf{d}|h_i)$ (popř. $g_i(\mathbf{d})$) určuje příslušnost testovaného vzorku k i -té třídě. Podobně tuto příslušnost určuje minimální hodnota $D_i^2(\mathbf{d})$. Tím je potvrzen hypotéza, že daný vzorek náleží do třídy i a ostatní hypotézy (o příslušnosti prvku k jiným třídám) jsou zamítnuty.

Ilustrujme popsany postup na konkrétním příkladě.

- Mějme testovací vzorek popsany dvěma formantovými kmitočty 700 Hz a 1000 Hz, to znamená $\mathbf{d} = [700 \ 1000]$.
- Na základě vztahu (12) vypočteme diskriminační funkce \mathbf{g}_i , $i = 1, 2, \dots, 5$ pro jednotlivé typy samohlásek, přičemž použijeme symbolické označení:
 $\mathbf{g}_A = -32, 16$; $\mathbf{g}_E = -60, 38$; $\mathbf{g}_I = -56, 30$; $\mathbf{g}_O = -20, 93$; $\mathbf{g}_U = -21, 72$.
- Na základě maximální hodnoty diskriminační funkce \mathbf{g}_O zařadíme neznámý testovaný vzorek \mathbf{d} do třídy samohlásek „O“. Tím je potvrzena hypotéza, že daný prvek představuje hlásku O a ostatní čtyři hypotézy jsou zamítnuty.
- Pro porovnání provedeme také klasifikaci na základě euklidovské vzdálenosti (1): $v_A = 587$, $v_E = 1018$, $v_I = 1152$, $v_O = 158$, $v_U = 151$. Minimální hodnotu má vzdálenost h_U , a proto zařazujeme neznámý testovaný vzorek \mathbf{d} do třídy samohlásek „U“.

D. Rozbor výsledků

Obrázek 2 nám zobrazuje rozdělení všech samohlásek z naší testovací databáze v rovině prvních dvou formantů F1 a F2. Plná čára elipsy zobrazuje vrstevnici ve vzdálenosti 2,45-násobku směrodatné odchylky od středních



Obrázek 2: Rozdělení českých samohlásek v rovině prvních dvou formantů F1 – F2

hodnot, což zahrnuje 95 % dat pro danou třídu⁵. Je zřejmé, že oblasti nejsou disjunktivní, a proto klasifikace s parametry v rovině F1-F2 nemusí být vždy úspěšná. Tento závěr potvrzuje i konfuzní tabulka, kterou získáme testováním na dané databázi. Provedeme-li uzavřené testování na této databázi, dostáváme výsledky uvedené v tabulce 2. Z této konfuzní tabulky můžeme zjistit, že ze 44 samohlásek „A“ bylo správně rozpoznáno 39 (což je 87 %). Naopak 2 samohlásky „A“ byly chybně klasifikovány jako „E“ a 3 samohlásky „U“ jako „O“.

Zvýšení úspěšnosti klasifikace lze dosáhnout zvýšením počtu parametrů a někdy rovněž volbou jiného typu parametrů, které vytvoří oblasti s menším překrýváním.

Kromě použitých formantových frekvencí jsou často používanými parametry keprální koeficienty. Tomuto způsobu parametrizace bude proto věnován následující příklad ilustrující možnosti bayesovské klasifikace frikativ.

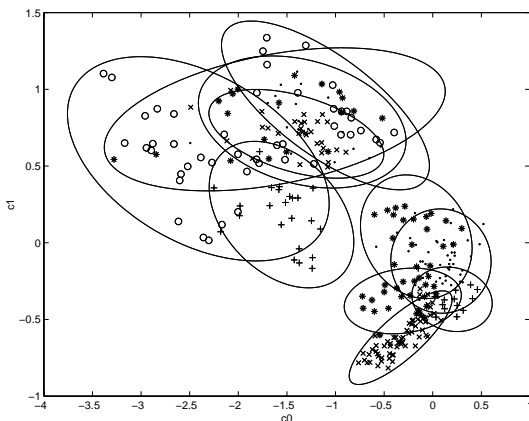
	A	E	I	O	U	%
A	39	2	0	3	0	87
E	2	36	6	0	0	82
I	0	9	32	0	0	78
O	0	0	0	36	3	92
U	0	0	0	4	27	87

Tabulka 2: Konfuzní tabulka českých samohlásek - první dva formanty - uzavřený test

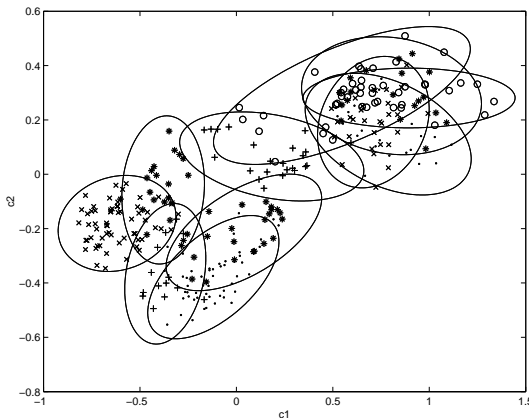
7.2. Příklad klasifikace frikativ

Frikativy (sonorní – R, L, J; šumové znělé – V, Z, Ž, H, Ř; šumové neznělé – F, S, Š, CH), podobně jako ostatní

⁵Data jsou převzata z práce [2].



Obrázek 3: Rozdělení českých frikativ v rovině keprstrálních koeficientů $c_0 - c_1$



Obrázek 4: Rozdělení českých frikativ v rovině keprstrálních koeficientů $c_1 - c_2$

souhlásky, nemají oproti samohláskám výraznou formantovou strukturu. Pro jejich popis však lze použít keprstrálních koeficientů⁶.

Na obr. 3 můžeme sledovat rozložení frikativ v rovině keprstrálních koeficientů c_0 a c_1 [12]. Obecně lze říci, že koeficient c_0 je úměrný logaritmu rozptylu signálu, koeficienty c_1 a c_2 popisují tvar vyhlazeného spektra signálu a vyšší keprstrální koeficienty popisují detaily ve spektru. Na obr. 4 je rozložení frikativ v rovině c_1 a c_2 .

Reálné keprstrum bylo počítáno ze segmentů signálu dlouhých 256 vzorků (16 ms při vzorkovací frekvenci 16 kHz) s překryvem 50%. Výsledky experimentu s klasifikací českých frikativ [12] v uzavřeném testu (data jsou od jednoho mluvčího) jsou uvedeny v tabulce 3.

Je zřejmé, že u některých hlásek opět dochází k překrývání oblastí, a proto mají nižší úspěšnost rozpoznání. Výsledky lze zlepšit, jak již bylo uvedeno, použitím většího počtu parametrů a především většího počtu trénovacích dat. Pro rozpoznávání řeči se typicky používá nejméně 10 parametrů včetně informace o jejich časovém vývoji. Ča-

F	H	J	L	R	Ř	S	Š	V	Z	Ž	∅
96	67	60	89	42	93	92	78	86	81	82	79
%	%	%	%	%	%	%	%	%	%	%	%

Tabulka 3: Rozpoznávací skóre pro česká frikativa pro první tři keprstrální koeficienty v uzavřeném testu pro jednoho mluvčího

sový vývoj popisují difference (prvního a druhého řádu) parametrů. Proto může být v jednom vektoru, popisujícího jeden segment řeči, i více než 30 parametrů. Rovněž frekvenční osa bývá nelineárně transformována, aby rozpoznávací systémy lépe aproximovaly fyziologické vlastnosti lidského ucha a tím dosáhly vyšší šumové odolnosti a spolehlivosti klasifikace. Modely tříd se trénují na podstatně větším množství dat, než jaké bylo použito v příkladech. To vše v uvedených ilustrativních příkladech nebylo použito, a proto jsou výsledky horší, než jakých dosahují rozpoznávače řeči. Struktur rozpoznávačů řeči je mnoho a jsou komplikované, nicméně nejlepších výsledků dosahují ty, které využívají bayesovský přístup.

Další příklad použití diskriminační analýzy pro klasifikaci řečových segmentů do fonetických tříd lze nalézt v práci [4]. Výsledky úspěšnosti klasifikace pomocí otevřeného testu byly následující: klasifikace do 3 tříd 96%, klasifikace do 8 tříd 73%, klasifikace do 20 tříd 44%. V prvních dvou případech lze tedy diskriminační analýzu úspěšně použít, pro velký počet tříd selhává.

8. Závěry

V článku byly popsány základy bayesovských technik a jejich použití pro klasifikaci dat. Na vybraných jednoduchých příkladech klasifikace řeči byl vysvětlen postup parametrizace dat a jejich následné klasifikace. Podrobně byly popsány vlastnosti a použití jednotlivých členů Bayesova vztahu. Především byl ilustrován vliv apriorní informace, která bayesovské techniky odlišuje od metody maximální věrohodnosti. Rovněž byl zmíněn význam evidence pro výběr modelu dat. Bylo ilustrováno, že bayesovské techniky jsou spolehlivější než metody založené na použití vzdálenosti v euklidovské metrice nebo metoda maximální věrohodnosti.

Poděkování

Tato práce byla podporována komplexním grantem „Hlasové technologie v podpoře informační společnosti“, GA 102/02/0124 a výzkumným záměrem „Transdisciplinární výzkum v oblasti biomedicínského inženýrství“, MSM 210000012.

Reference

- [1] Anděl, J.: *Statistické metody*. Matfyzpress, Praha, 1998.

⁶Více o keprstrální analýze lze nalézt např. v [16], [14].

- [2] Čmejla, R.: *Rozpoznávání explozí v řečovém signálu*, kandidátská disertační práce, CTU, Praha, 1993.
- [3] Čmejla, R., Sovka, P.: *Application of Bayesian Autoregressive Detector for Speech Segmentation*, International Conference on Signal Processing Applications and Technology Proceedings, Miller Freeman Inc., California, USA, 1999.
- [4] Čmejla, R., Sovka, P.: *Blind Location of Phonetic Boundaries*, Proceedings of 17-th International Congress on Acoustics, September 2-7, Rome, Italy, 2001
- [5] Čmejla, R., Sovka, P.: *Estimation of Boundaries between Speech Units Using Bayesian Changepoint Detectors*. Text, Speech and Dialog, 4th International Conference, (Lecture notes in computer science; Vol. 2166), Springer-Verlag, Berlin - Heidelberg - New York, Germany, pp.291-297, 2001.
- [6] Čmejla, R.: *Bayesovská detekce náhlých změn*. Habilitační práce, FEL ČVUT, Praha, 2001.
- [7] Harrington, J., Cassidy, S.: *Techniques in speech acoustics*. Kluwer Academic Publishers, Dordrecht-Boston- London, 1999.
- [8] Kotek, Z., Mařík, V., Hlaváč, V., Psutka, J., Zdráhal, Z.: *Metody rozpoznávání a jejich aplikace*. Academia, Praha, 1993.
- [9] Králík, J., Čmejla, R., Sovka, P., Stančák, A.: *Head Reconstruction and Localisation of Brain Activity using Bayesian Evidence*, Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Istanbul, Turkey, 2001.
- [10] Meloun, M., Militký, J.: *Statistické zpracování experimentálních dat*. Ars magna, Praha 1998.
- [11] Novovičová, J.: *Pravděpodobnost a matematická statistika*. Skripta FD ČVUT, Praha, 1999.
- [12] Pech, J.: *Klasifikace frikativ*. Semestrální projekt v předmětu Digitální zpracování řeči, Praha, 2000.
- [13] Psutka, J.: *Komunikace s počítačem mluvenou řečí*. Academia, Praha, 1995.
- [14] Sovka, P., Pollák, P.: *Vybrané metody číslicového zpracování signálů*. Ediční středisko ČVUT, Praha, 2001.
- [15] Vaseghi, S.V.: *Advanced Signal Processing and Digital Noise Reduction*, Wiley – Teubner, Chichester, 1996.
- [16] Vích, R.: *Z Transform Theory and Applications*, D. Reidel Publishing Company, Dordrecht, 1987.

Clinical interpretation of the acoustic pressure measurement in the human ear canal

Jan Mejzlík*, Zdeněk Škvor**, Libor Husník**, František Rund**

*ORL oddělení, Nemocnice Pardubice, Kyjevská 44, 532 03 Pardubice

**ČVUT, Fakulta elektrotechnická, katedra radioelektroniky, Technická 2, 166 27 Praha 6

Na dvou spánkových kostech kadaverózních dárců byly mikrofonními sondami měřeny přenosové charakteristiky akustického tlaku zevního zvukovodu a bubínku. Výsledky měření prokázaly, že zevní zvukovod výrazněji ovlivňuje pouze akustické vlny na frekvencích nad 1,5 kHz. Zvuky na frekvencích nad 4 kHz jsou též významně ovlivněny bubínkem.

1. Introduction

The sound transfer functions (STF's) of the external ear canal in homogenous acoustic field are known [5], [6], [7].

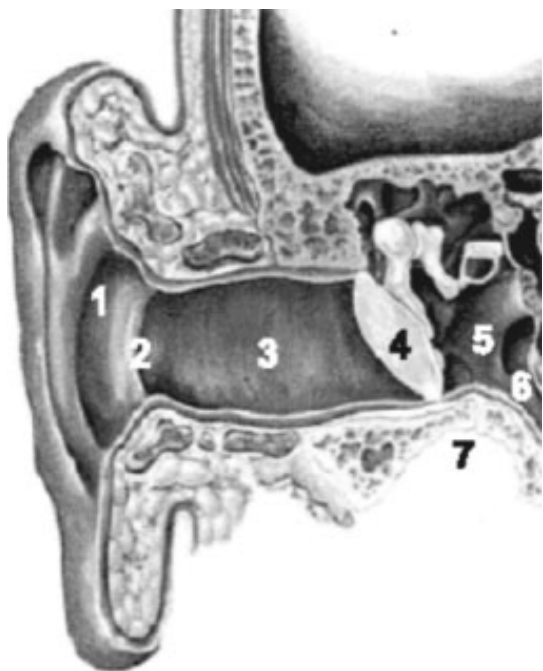


Fig. 1: Cross-section of the external and middle ears: 1. cavum conchae, 2. meatus acusticus externus, 3. ear canal, 4. tympanic membrane, 5. tympanic cavity with middle ear bones, 6. Eustachian tube, 7. temporomandibular joint.

To perform the measurement in the lumen of an ear canal the microphone probes are usually used [5]. The authors suggest, that the alien material of the probe microphones influence acquired values. The ear canal is a tube between the meatus acusticus externus and the tympanic membrane (TM). The meatus acusticus externus is located in the cavum conchae (Fig.1). The cross-section of the external ear canal is in the form of an ellipse the longer axis of which has near the ear drum a horizontal orientation.

The longer axis of the ellipse has near the entrance to the ear canal almost vertical orientation and its length is approximately 9 mm. When approaching the tympanic membrane the axis bends downwards, its orientation becomes horizontal and its length is 8 mm. The shorter axis is around 6,5 mm in length near the ear canal entrance and 4,5 mm on average near the tympanic membrane. The smallest lumen of the ear canal is isthmus, which is located 20 mm from the concha floor. Its longer axis is on average 8 mm in length, its shorter axis about 6 mm in length. The length of the frontal ear canal wall measured from the frontal edge of tragus to the frontmost edge of the ear drum is on average 35 mm, the length of the back wall, measured from the cavum conchae, is 25 mm. Lengthwise axes of both ear canals converge at the angle 80°, open in the dorsal plane.

Skin of the cartilaginous portion of the meatus is covered by numerous hairs, which are bent laterally and thicken in the course of time. There is also a number glands in the skin, ceruminous glands, which in the depth of the skin form a compact layer. These glands can also be found in a small area of the bony portion of the canal. Glands produce a golden brown wax, cerumen.

2. Aims

The aim of this study was to assess how the measured STF's of the cadaverous auditory meatus correspond with the curves presented in literature and if these characteristics are distorted by interference of waves heading to the tympanic membrane and the waves reflected from the tympanic membrane. The STF of the tympanic membrane was also measured, as the vibrations of the ear drum depend not only on acoustic pressure on the outer surface of the membrane but also on pressure on its inner surface, hence on their difference [5].

3. Material and methods

Two cadaverous temporal bones preserved in 10 % solution of formalin were used for the study. The auricle was not object of the research and therefore was removed from

the temporal bones covered by soft tissues. The ear canal ended laterally by a sharp margin of the skin in the cavum conchae.

The middle ear was reached by cutting off the top of the temporal bone pyramid in the plane parallel with the ear drum in the distance of 20 mm. This ensured the access to the middle ear through the Eustachian tube.

Malleus, incus and stapes as well as middle ear cavities were preserved in a way, which did not change their mechanical properties. Specimen was fastened by a screw through a hole with a corresponding size drilled into the lower part of the temporal bone to the metal stand consisting of a steel pole with the diameter of 10 mm and a thread of 6 mm.

A set of two probes [3] was placed on other supports. Probes were inserted through the temporomandibular joint and into the middle ear through the dilated Eustachian tube.

Different probe locations (Fig. 2.) chosen to obtain the acoustic pressure transfer function between the reference probe (RS) and the measuring probe (MS) result in following transfer characteristics depicted in Fig. 3–5.

To avoid penetration of the sound along the probe and to ensure the measured data represented sound pressure levels inside ear canal and middle ear, the place where the probe was inserted into the tissues was sealed by silicon vaseline as was the residual hole. Three measurements with different location of probes (Fig. 2) were performed on each of the temporal bones under stable laboratory conditions, temperature, pressure, humidity using the same instruments. In evaluation of measured data we focused on finding maxima and minima of acoustic pressure. The acoustic pressure frequency characteristics as well as the rate of relative increase or reduction is one of standards used in assessment of the acoustic pressure spectrum measured in the ear canal, middle and inner ears.

4. Results

The STF's for the three different microphone probe placements in the right temporal bone are depicted in Fig. 3–5.

5. Discussion

When evaluating the measured curves we have to realize that the ear drum has three basic functions: transmission of the sound to the middle and inner ears, amplification, and shading of the round window.

The tympanic membrane protects the round window (fenestra rotunda) by absorbing and reflecting part of the acoustic energy propagating in the ear canal from noise which might provoke interference of acoustic waves in the inner ear. Intact tympanic membrane represents an obstacle especially for noise at frequencies higher than 4 kHz.

Low levels of acoustic pressure at high frequencies behind the ear drum may partially explain very small al-

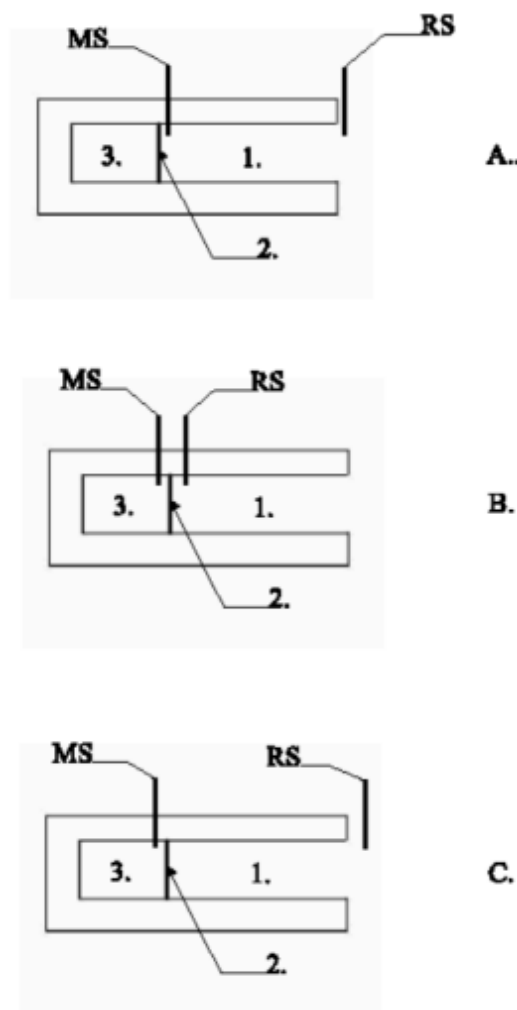


Fig. 2: Positions of the microphone probes (MS, RS) for the three types of measurements: A–Transfer characteristics of the ear canal. B–Transfer characteristics of the eardrum. C–Transfer characteristics of the ear canal and the eardrum. (1. ear canal, 2. tympanic membrane, 3. tympanic cavity)

terations of hearing threshold at high frequencies in noise induced hearing loss (NIHL), meanwhile maximal acoustic pressure in the range from 2 to 4 kHz respond to maximal acoustic pressure of hearing changes in NIHL.

The acoustic pressure maximum (3,2 kHz) in the ear canal correspond with one of the minima in the tympanic cavity. It should explain the greater movement of the tympanic membrane in this range, as it is determined by the difference of the pressure in front of and behind the tympanic membrane. In this case we can also suggest that frequency characteristics of acoustic pressure behind the eardrum is greatly influenced by the middle ear cavity [4].

According to other authors [1] the increase of acoustic pressure in the ear canal is canceled in the next passage through the middle ear by transfer characteristics of middle ear ossicula.

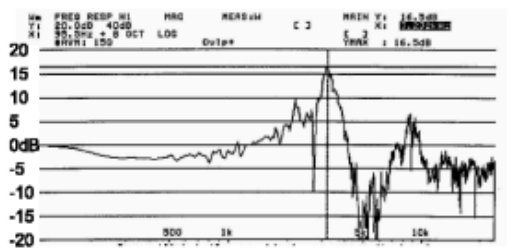


Fig. 3: Sound transfer function – Ear canal

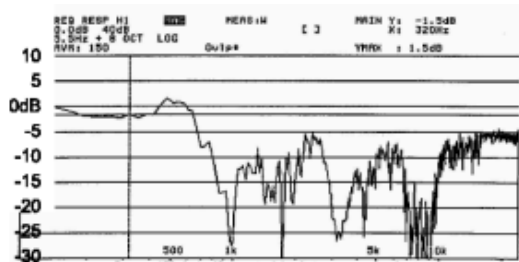


Fig. 4: Sound transfer function – Tympanic membrane

Interesting results were found in so-called manubriostapedopexis. Manubrio-stapedopexis is a kind of a middle ear operation when the handle (manubrium) of the hammer (malleus) is fixed to the stapes and anvil (incus) is missing. The air-bone difference in the ideal case is 10 dB in the whole hearing range [2].

The above mentioned facts give a rise to a question how much acoustic waves on the round window influence transfer of the sound to the inner ear with the intact tympanic membrane and middle ear ossicles. The acoustic pressure changes phase during the passage through the ear drum at some frequencies. And furthermore there is a resonant increase of the waves in the middle ear.

6. Conclusion

We can state three basic findings according to our results

1. Transfer characteristics of acoustic pressures acquired in front of the tympanic membrane correlate with characteristics measured behind the tympanic membrane only in the frequency range from 0,5 to 4 kHz.
2. Acoustic pressure at frequencies lower than 1,5 kHz is very little influenced by acoustic properties of the ear canal.
3. Acoustic pressure transfer function of the tympanic membrane at frequencies higher than 4 kHz cannot be reliably evaluated with this type of measurement.

Acknowledgement

This work has been supported by the research grant of Ministry of Health NK6179-3.

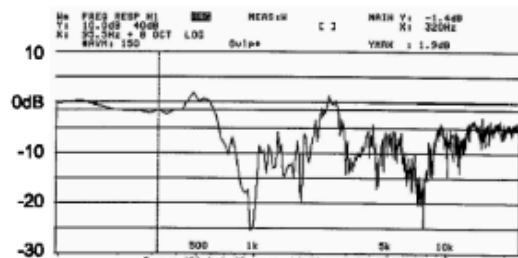


Fig. 5: Sound transfer function – Ear canal and tympanic membrane

References

- [1] Bornitz, M., Zahnert, T., Hardtke, H., Hüttenbrink, K. B.: Identification of parameters for the middle ear model, *Audiol. Neurootol.*, 3-4(4), p. 163–169, 1999.
- [2] Hroboň, M., Betka, J., Skřivan, J.: Sluchové zisky u manubrio-stapedopexie, *Otorinolaryngol.*, 3(44), p. 159–161, 1995.
- [3] Husník, L., Mejzlík, J., Rund, F., Juřica, M., Michna, I., Gajdůšek, M., Škvor, Z.: Measuring Apparatus for Determination of Acoustic Pressure in the Human Ear Canal, *Akustické listy*, 3(7), p. 7–9, 2001.
- [4] Pellant, K., Mejzlík J., Šremr J., Pellant A.: Transmission Properties of Outer Ear Canal, *Proceedings of 3rd International Conference Mechatronics, Robotics and Biomechanics*, p. 245–252, Třešť, 2001.
- [5] Shaw, E.A.G.: The External Ear. In.: Keidel, W.D. and Neff, N.P.(eds.), *Handbook of Sensory Physiology*, Vol. 5/1, p.455–490, Berlin: Springer Verlag, 1974., p.455–490.
- [6] Schuknecht, H.F.: *Pathology of the Ear*, 2nd edition: Lea & Fibiger, Malvern, 1993, p. 289-294.
- [7] Syka, J., Voldřich, L., Vrabec, F.: *Fyziologie a patologie zraku a sluchu*. 1. vyd., Avicenum, Praha, 1981, p. 176-177.

Pokyny pro přípravu příspěvků do Akustických listů

Příspěvky do Akustických listů je možné napsat v jazyce českém, slovenském nebo anglickém. Příspěvky se dodávají v elektronické podobě na e-mailovou adresu *brothan@fel.cvut.cz* nebo *jiricek@fel.cvut.cz*.

Pro vypracování příspěvku je možné použít textový editor:

- **L^AT_EX** (do něj jsou příspěvky převáděny)
- **Word**
- **jiný** – zvlášť text a obrázky. Doporučujeme předem konzultovat s redakcí.

Na webové stránce

http://www.czakustika.cz/czas_cz.htm

jsou umístěny ukázky příspěvků pro L^AT_EX, resp. Word včetně použitého classu, resp. šablony. Jejich použití velice zjednoduší tvorbu vlastního příspěvku.

redakce

Akustické listy: ročník 8, číslo 2 červen 2002
Vydavatel: Česká akustická společnost, Technická 2, 166 27 Praha 6
Počet stran: 16 Počet výtisků: 200

ISSN: 1212-4702

Vytisklo: Ediční středisko ČVUT

Číslo připravili: Marek Brothánek, Ondřej Jiříček, Jan Kozák

© ČsAS

Příspěvky nejsou redakčně upravovány. Za jazykovou úpravu odpovídají jejich autoři.

Uzávěrka příštího čísla Akustických listů je 13. září 2002.

NEPRODEJNÉ!