

Klasifikace Landau-Kleffnerova syndromu malých dětí

Petr Zlatník

ČVUT FEL, K13131 Katedra teorie obvodů

1. Abstrakt

Tento příspěvek pojednává o klasifikaci *Landau-Kleffnerova syndromu*, který se projevuje u malých dětí. Daný problém je řešen ve spolupráci s Fakultní nemocnicí v Motole, kde jsou nahrávány řečové promluvy od postižených dětí. Syndrom lze detekovat a klasifikovat jednak ze signálu mozkové aktivity (EEG) nebo testováním promluv. Zde je využíváno testování a porovnávání promluv s promluvami od zdravých dětí pomocí principů rozpoznávačů řeči. Do budoucna by měla vzniknout metoda, která bude schopna jednak porovnáním se zdravými dětmi určit jak hodně je dané dítě postižené a měla by být schopna rozlišit, zda se syndrom u postiženého dítěte léčbou lepší či ne. Uvedené algoritmy jsou naprogramovány v Matlabu verze 6.0.0.88. Pro práci s databázema řečových promluv byl využit Cool Edit 2000.

2. Úvod

Landau-Kleffnerův syndrom je vzácná nemoc, která se projevuje zřídka u malých dětí mezi třetím až sedmým rokem. Jedná se o náhlou nebo postupnou poruchu té části mozku, která zodpovídá za porozumění a výslovnost řeči. Daný syndrom se tedy projevuje abnormálními stavy v EEG signálu. Pomocí tohoto signálu je skutečně ve většině případů tento syndrom lékaři sledován a klasifikován.

Typicky se dané dítě nejprve vyvíjí normálně a okolo třetího roku se začne daný syndrom projevovat neschopností dítěte přesně zopakovat řečené slovo a větu, má problémy s výslovností a může mít problémy s učením. V extrémních případech dítě nerozumí ani svému vlastnímu jménu a tím se jeví na první pohled jako sluchově postižené, není schopné rozeznat okolní zvuky, jako je telefon. S tím souvisí stavy depresí, podrážděnosti a neschopnosti se soustředit.

3. Použité metody klasifikace

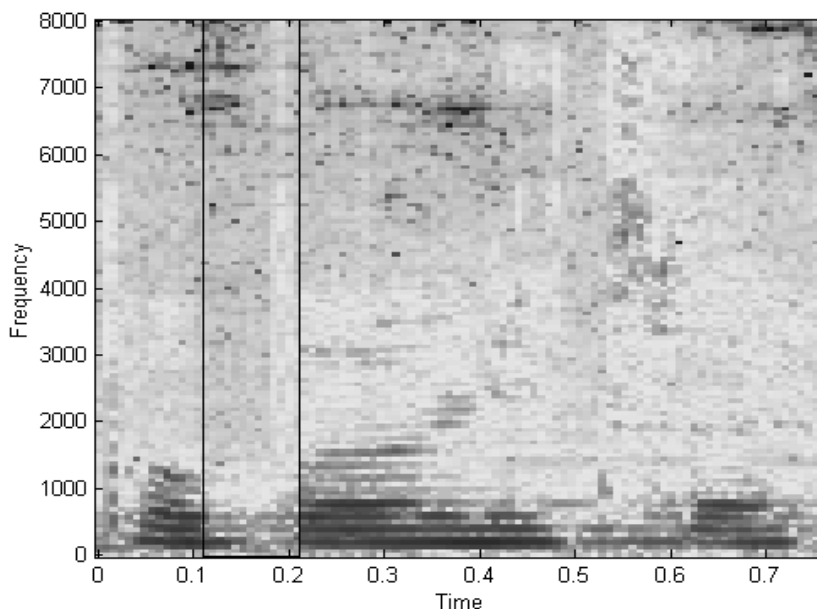
3.1. Porovnání délek okluze ve slovech

První pokusy při klasifikaci Landau-Kleffnerova syndromu vedly k tomu, že byl předpoklad větší délky *okluze* vybraných slov nemocných dětí oproti zdravým. Okluze je speciální typ souhlásky. Souhláska obecně je vytvářena vzduchovou turbulencí vznikající třením proudu vzduchu při výdechu o překážku, kterou tvoří artikulační orgány jako například rty, zuby nebo špička jazyka.

Pokud je překážka úplná (nazývá se také závěr), vzniká v okamžiku jejího přerušení krátký šum, který se podobá výbuchu, explozi. Souhlásky, které vznikají při využití úplné překážky se nazývají *okluze*. Mezi ně patří

m, n, ň, g, d, d', b, p, t, t', k.

Protože nemocné děti mají problémy s vyslovením slova a jejich reakce při opakování daného slova je většinou pomalejší než u zdravých dětí, byl předpoklad, že bude větší délka okluze jednotlivých hlásek ve vybraných slovech než je tomu u zdravých dětí. Daná myšlenka byla testována na slově 'papír', kde se testovala ve spektrogramu délka okluze 'p'. Průběh je na obrázku 1, kde je daná délka na časové ose vyznačena dvěma svislými čarami. Průběh spektrogramu je z promluvy zdravého dítěte.



Obrázek 1: Spektrogram slova 'papír' s vyznačením délky okluze

Bohužel délka okluzí nemocných dětí není větší v porovnání se zdravými. Z poslechu jednotlivých promluv bylo zjištěno, že to je dáno tím, že nemocné děti řeknou hlásku 'p' stejně rychle jako zdravé, ale už nejsou schopné vyslovit správně hlásku 'r' na konci slova.

Danou metodu proto nelze využít.

3.2. Využití algoritmu borcení časové osy (DTW)

Další možnost pro klasifikaci syndromu bylo využití algoritmu *borcení časové osy* (DTW, Dynamic Time Warping). Tento algoritmus byl popsán v mnoha publikacích, proto je zde o něm zmíněno stručně, pro tyto účely se vycházelo z [3]. Nejprve bylo nutno natrénovat průměrné modely jednotlivých referenčních slov z promluv od co nejvíce zdravých dětí, bylo jich k dispozici přibližně třicet. Toho se dosáhlo tímto způsobem. Jednotlivé promluvy byly nasegmentovány na časové úseky 10 *ms*. Každý ze segmentů byl popsán jedenácti LPC koeficienty pomocí Burgova algoritmu, které byly dále převedeny na kepstrální. Protože každé dítě namluví stejné slovo různě rychle, tak se samozřejmě každý model referenčního slova skládal z různého počtu segmentů při dodržení kroku segmentace 10 *ms*. Proto bylo nutné normovat každý model na průměrný počet segmentů, aby šel vypočítat průměrný model každého slova ze všech dětí. Toho se dosáhlo pomocí vztahů v [3]. Nakonec byl vypočítán průměrný model. Stejným způsobem bylo popsáno slovo od nemocného dítěte včetně výše popsaného průměrování. Potom již bylo možné vypočítat matici kepstrálních vzdáleností, výpočet byl proveden pomocí klasických Euklidovských vzdáleností. Tato matice byla procházena ve směru nejmenší vzdálenosti z jednoho rohu do protějšího (nebývá směr po úhlopříčce, ale po obecné křivce), jak je naznačeno na obrázku 2 a byl vypočítán *naakumulovaný součet minimálních kepstrálních vzdáleností S*, který lze matematicky popsat jako

$$S = S + \min(S1, S2, S3), \text{ kde}$$

$$S1 = S + d(i + 1, j)$$

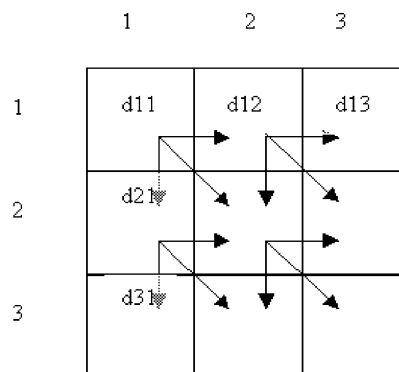
$$S2 = S + d(i, j)$$

$$S3 = S + d(i, j + 1), \text{ kde}$$

i – je číslo sloupce,

j – je číslo řádku matice

kepstrálních vzdáleností.



Obrázek 2: Průchod matice kepstrálních vzdáleností

Tento výpočet se provádí v cyklu, tedy po postoupení na další políčko se provádí znovu výpočet součtu kepstrálních vzdáleností pro 3 nejbližší okolní políčka (*S1, S2, S3*) až do průchodu celou maticí, tzn. od roku 1,1 do rohu *M, N*, kde *M, N* jsou počty segmentů referenčního a srovnávaného signálu. V mém případě samozřejmě *M = N*, protože je jak průměrný model tak i testované slovo normováno na průměrný počet segmentů. Byl

předpoklad, že když se budou s průměrným modelem porovnávat jednotlivá slova od zdravých dětí, tak v matici kepstrálních vzdáleností půjde vymezit oblast zdravých dětí a testovat, zda nemocné dítě spadne dovnitř nebo ne. Pokud spadne vně, tak se jedná o nemocné dítě. Bohužel se tato domněnka nepotvrdila. Při porovnání nemocného dítěte s průměrným modelem docházelo k tomu, že průchod maticí byl více po úhlopříčce než při porovnání zdravého dítěte.

Při porovnávání nemocných dětí s průměrným modelem však ve většině případů byla podstatně větší *celková naakumulovaná vzdálenost* S než při porovnávání se zdravými. Proto se bylo ubíráno tímto směrem. Metoda ale nebyla stoprocentně korektní. Algoritmus DTW vykazoval neschopnost rozpoznání například nevyslovení jedné hlásky ve slově. Když zdravé dítě vyslovilo slovo 'mateřídouška' pomaleji než většina dětí z kterých byl počítán průměrný model, tak celková naakumulovaná vzdálenost byla vysoká a algoritmus označil zdravé dítě chybně jako nemocné. Pokud ale nemocné dítě vyslovilo 'matouška' místo 'mateřídouška', docházelo k tomu, že vzdálenost byla nízká a algoritmus označil chybně naopak nemocné dítě za zdravé.

Protože ani tato metoda není příliš spolehlivá, byla vyzkoušena metoda založená na trochu jiném principu.

3.3. Testování odchylek kepstrálních koeficientů

Jak již bylo uvedeno výše, algoritmus DTW není spolehlivý. Proto byla vyzkoušena ještě jedna možnost klasifikace vycházející ze základu algoritmu DTW. Jednotlivé promluvy od zdravých i nemocných dětí byly opět segmentovány s délkou časového úseku 10 *ms*. Každý segment byl popsán jedenácti kepstrálními koeficienty a bylo provedeno normování na průměrný počet segmentů pro jednotlivá slova a mluvčí.

Pro testované slovo se z promluv všech zdravých dětí určila nejmenší a největší hodnota každého kepstrálního koeficientu ve všech segmentech a tím vznikly intervaly koeficientů zdravých dětí. Potom bylo porovnáváno, kolik hodnot kepstrálních koeficientů nemocného dítěte přes všechny segmenty spadne mimo intervaly definované zdravými dětmi a kolik dovnitř. Počet hodnot koeficientů padajících mimo je úměrný stupni postižení daného dítěte. Ani tento algoritmus však není schopen spolehlivě rozlišit zdravé dítě od nemocného. Problém je opět v tom, jako u DTW, že pokud většina dětí trénovací databáze namluví určité slovo přibližně stejně rychle, pak při klasifikaci slova od dítěte které ho namluví pomaleji dojde k označení zdravého dítěte za nemocné. Naopak pokud nemocné dítě vypustí ve slově hlásku, občas dochází k tomu, že to algoritmus nezaznamená a naopak klasifikuje nemocné dítě za zdravé.

4. Zhodnocení

Na závěr lze konstatovat, že srovnávání délek okluze u zdravých a nemocných dětí nelze

využít. Neprokázalo se, že by byla u nemocných dětí větší délka. Bylo zjištěno, že její délka závisí především na tom, jak rychle dané dítě testované slovo vysloví.

Algoritmus DTW je nespolehlivý, pro některá slova více, pro některá méně, neuvažuje se další jeho využití v této podobě pro další práci.

Testování odchylek kepstrálních koeficientů opět není spolehlivé. Daný algoritmus byl testován již na rozšířené databázi, jak na delších slovech jako je 'mateřídouška', 'televize', tak i na kratších slovech která se obtížně vyslovují jako je 'krk', ale i na jednotlivých souhláskách a samohláskách. Pro další práci by bylo vhodné rozšířit databázi jak zdravých tak i nemocných dětí a algoritmus otesovat znovu.

Srovnání úspěšnosti rozlišení všech tří algoritmů zdravých dětí od nemocných v % je v tabulce 1.

SLOVO	*1	*2	*3
KRK	-	96	83
MOTOVIDLO	-	83	66
TELEVIZE	-	97	83
POPELNICE	35	83	66
MATEŘÍDOUŠKA	-	67	66
RŮZNOBAREVNÝ	-	67	66

Tabulka 1: Tabulka úspěšnosti v % všech tří algoritmů

pozn.: *1 - Úspěšnost srovnávání délek okluze v % pro slovo 'popelnice'

pozn.: *2 - Úspěšnost algoritmu DTW pro různá slova v %

pozn.: *3 - Úspěšnost testování odchylek kepstrálních koeficientů pro různá slova v %

5. Výhled pro další práci

Ve výhledu je počítáno s tím, že bude využit rozpoznávač na principu skrytých Markovových modelů (HMM) s využitím softwaru HTK. Zde je předpoklad, že narozdíl od výše popisovaných algoritmů bude schopen lépe rozlišovat a popisovat jednotlivé hlásky ve slovech a nebude docházet tolik často k tomu, že bude označeno zdravé dítě za nemocné a naopak.

6. Poděkování

Tato práce je podporována grantem č. 13131/13/03016 "Modelování biologických a řečových signálů" a výzkumným záměrem č. 102/03/H085.

7. Literatura

- [1] P. Sovka, P. Pollák : *Vybrané metody číslicového zpracování signálů*. Ediční středisko ČVUT Praha, 2001
- [2] V. Davídek, P. Sovka : *Číslicové zpracování signálů a implementace*. Vydavatelství ČVUT Praha, 2002
- [3] J. Psutka: *Komunikace s počítačem mluvenou řečí*. Vydala Academia Praha, 1995, tisk CENTA, spol. s. r. o., Veverí 39, Brno

8. Kontaktní informace

Ing. Petr Zlatník
České vysoké učení technické v Praze
ČVUT FEL K13131, Technická 2, 166 27 Praha 6, Česká republika
Tel.: (+420) 22435 2820
E-mail: zlatnik.petr@seznam.cz