

PARAMETRIZACE PRŮBĚHU KŘIVKY PRO HODNOCENÍ CHRAPTIVOSTI

A. Stráník, R. Čmejla

Katedra teorie obvodů, Fakulta elektrotechnická, ČVUT v Praze

Abstrakt

Chraptivost je obecné označení pro přítomnost patologických příměsí v hlase a je také jedním z prvních příznaků onemocnění hlasového ústrojí. Pro hodnocení chraptivosti se často používají různé parametrizace, které kvantifikují např. míru šumové příměsí v hlase. V současné době jsou nejčastěji k popisu použity pouze střední hodnoty nebo směrodatné odchylky takových parametrů, které ovšem nezohledňují rozdílný způsob nakmitávání a dokmitávání hlasivek na začátku a na konci fonace u zdravých a chraptivých hlasů. V příspěvku je ověřena vhodnost použití parametrů CL (Curve Length) a LRE (Line Regression Error) pro detailnější popis časového průběhu parametrů při hodnocení chraptivosti.

1 Úvod

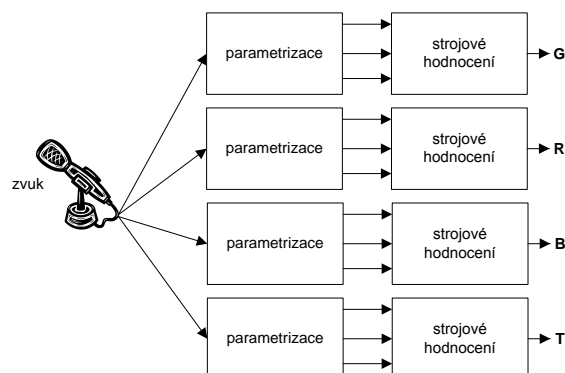
Chraptivost je obecné označení pro přítomnost patologických příměsí v hlase a je také jedním z prvních příznaků onemocnění hlasového ústrojí. Je provázena různými jevy jako například chraptivostí, chraplavostí, dyšností nebo hlasovým napětím, které při běžné komunikaci vnímá nejen školovaný specialista, ale i běžný laik. Tyto jevy nemusí být pro posluchače vždy nepříjemné, nicméně, jak je zmíněno výše, mohou být předzvěstí počínajícího vážnějšího onemocnění hlasového ústrojí.

Pro určení míry chraptivosti odborná veřejnost vypracovala několik postupů, jak od sebe rozlišovat různé druhy a tíže chraptivosti. Tyto metody jsou jak subjektivní, tak objektivní. Mezi subjektivní metody se řadí různé poslechové testy, při kterých odborník poslouchá hlasový projev pacienta a na základě vlastních zkušeností vyplňuje určité formuláře ve kterých je popsáno, na jaké aspekty řečového projevu by se měl zaměřovat. Objektivních metod existuje celá řada, některé jsou invazivní (laryngoskopie, laryngostroboskopie, . . .) a jiné neinvazivní (elektroglotografie, analýza akustického signálu). Z hlediska ceny vyšetření a komfortu z pohledu pacienta je vyšetření pomocí analýzy akustického signálu nejpříjemnější. Metody akustické analýzy hlasu mají také výhodu v tom, že zpětně neovlivňují pacienta, jako je tomu například u laryngoskopie.

Akustická analýza lidského hlasu je poměrně dlouho řešený problém, lze například nalézt odkazy na články ze sedmdesátých let, které se tímto fenoménem zabývají. To ovšem neznamená, že by téma bylo vyčerpané, ba právě naopak – akustické analýzy chraptivých hlasů jsou stále velmi aktuálním tématem.

Přítomnost chraptivosti lze popsat celou řadou parametrů, jako například HNR (*Harmonics-to-Noise Ratio*), CPP (*Cespral Peak Prominence*), GNE (*Glottal-to-Noise-Excitation Ratio*) apod. Ve studiích zabývajících se popisem chraptivosti je často používána pouze střední hodnota μ a směrodatná odchylka σ těchto parametrů [1, 2]. Zajímavou informaci, která by mohla sloužit k lepšímu popisu ovšem může nést i časový průběh výše zmíněných parametrů a to především kvůli rozdílnému nakmitávání a dokmitávání hlasivek na začátku a na konci prodloužené fonace.

Naším cílem je vyvinout automatický systém, který by odborníkům umožnil objektivizovat míru chraptivosti v hlase, viz Obr. 1.



Obrázek 1: Vizualizace funkce systému zamýšleného pro objektivní hodnocení chraptivosti.

2 Databáze

K experimentům byla využita databáze nahrávek Foniatrické kliniky 1. LF UK a VFN, která vznikala v letech 1977 až 1997 jako součást vyšetření. Nahrávky byly pořizovány v odhlučněné místnosti na kotoučové magnetofony. Záznamy byly v roce 2010 digitalizovány externí firmou s vzorkovací frekvencí 44,1 kHz a rozlišením 16 bitů na vzorek.

Nahrávky obsahují standardní vyšetření, mimo jiné čtení samohlásek /a/, /e/, /i/, /i/, /u/ a standardní text „Podzim na starém bělidle“. V současné době analyzujeme pouze fonaci vokálu /a/.

Pět logopedů z Foniatrické kliniky 1. LF UK a VFN hodnotilo modifikovanou stupnicí GRBAS (Grade, Roughness, Breathiness, Asthenia, Strain - celkový dojem, chraptivost, dyšnost, slabost, namáhání) celkem 469 promluv (72 zdravých a 397 chraptivých). Modifikace stupnice spočívá ve „sloučení“ dvou nejhůře hodnotitých parametrů¹ A a S do parametru T (Tension), který označuje napětí v hlase.

V použité stupnici GRBT používají hodnotitelé následující stupně: pro parametry G, R, B hodnotí v rozpětí 0 (nejlepší/nejméně přítomný) až 3 (nejvíce přítomný) body, parametr T je hodnocen v rozsahu -2 (největší hlasová slabost) až +2 (největší přemáhání hlasu) body.

Každá nahrávka byla odborníky hodnocena dvakrát, vždy s odstupem minimálně dvou týdnů. Odborníci při hodnocení znali pouze pořadové číslo nahrávky, které bylo při opakovaném hodnocení změněno.

3 Metodika

Akustické parametry pro hodnocení chraptivosti byly implementovány v prostředí MATLAB a to především z důvodu snadnější pozdější implementace do programovacího jazyku C#. Pro hodnocení chraptivosti bylo použito jak klasických tak původních parametrů.

3.1 Parametrizace akustického signálu

Mezi běžně používané parametry při hodnocení chraptivého hlasu patří HNR (Harmonics-to-Noise Ratio), např. [6, 7]. Tento parametr kvantifikuje míru šumové příměsi ve znělých úsecích řeči. Dyšní pacienti dosahují výrazně nižších hodnot HNR než pacienti zdraví. HNR lze odhadovat jak v oblasti časové [8], frekvenční [6, 9] i cepstrální [10]. K odkadu lze také využít waveletů [11].

Parametr CPP (Cepstral Peak Prominence) je dalším často používaným parametrem. Vyjadřuje míru harmonického buzení hlasového signálu a je určen jako rozdíl mezi amplitu-

¹Analýzy spolehlivosti subjektivního hodnocení byly provedeny např. ve studiích [3, 4, 5].

dou maxima v cepstru signálu, které odpovídá buzení a hodnotou aproximační přímky cepstra. U dobře znělých hlasů dosahuje výrazně vyšších hodnot než u hlasů dyšných nebo chraplavých.

GNE (Glottal-to-Noise Excitation Ratio) patří do skupiny poměrně nových parametrů, které popisují míru vybuzení akustického signálu hlasivkami. Autoři při odvození parametru vycházeli z faktu, že dobrý hlasivkový pulz lze po inverzní filtraci signálu poměrně dobře nahradit Diracovým pulzem, který rovnoměrně vybudí frekvenční spektrum v celém rozsahu. GNE pak popisuje míru korelace obálek inverzně filtrovaného signálu filtrovaného v různých frekvenčních pásmech [12, 13].

Míru periodicity signálu lze popsat i pomocí maxima autokorelační funkce v určitém časovém pásmu, které odpovídá pravděpodobné hlasivkové frekvenci (viz [1], parametr RPK), nebo pomocí detekce maxima, které skutečně odpovídá hlasivkové frekvenci (parametr SPS – Signal Periodicity Strength).

Parametr HLR (High-to-Low Energy Ratio) zohledňuje fakt, že energie řečového signálu je převážně na nižších frekvencích [12]. HLR vyjadřuje poměr energie signálu v pásmu 0-4 kHz ku energii signálu v pásmu 4 kHz až $f_s/2$.

HT a H1A je dvojice parametrů, která popisuje vzájemný vztah 1. a 2. harmonické. Zatlímco HT udává směrnici spojnice 1. a 2. harmonické, parametr H1A udává poměr jejich amplitud [1].

Další parametry použité se častěji používají při hodnocení a klasifikaci frikativních hlásek /s/, /š/, /z/, /ž/. Mezi tyto parametry patří SR (Spectral Rolloff), SPECW (Spectral Width) a SPECD (Spectral Deformation) [14].

Hodnota parametru SR určuje, pod jakou frekvencí leží XY procent celkové energie signálu. Např. značení SR30 vyjadřuje frekvenci, pod kterou leží 30% celkové energie signálu. Při použití tohoto parametru vycházíme z předpokladu, že převážná část energie leží na nižších frekvencích – dyšné hlasy ovšem mají šumovou příměs i na vyšších frekvencích a tím pádem bude hodnota parametru SR vyšší.

Parametry SPECD a SPECW pochází z hodnocení elektromyografických signálů a dávají do poměrů spektrální momenty. Definovány jsou následovně:

$$\text{SPECD} = \frac{\sqrt{\frac{M_2}{M_0}}}{\frac{M_1}{M_0}}, \quad \text{SPECW} = \sqrt{\frac{M_2}{M_1} - \left(\frac{M_1}{M_0}\right)^2},$$

kde $M_n = \sum_i P_i f_i^n$, přičemž P_i je spektrální výkonová hustota na frekvenci f_i .

Dále jsou použity odhady fundamentální frekvence, které získáme při výpočtu parametrů HNR, CPP a SPS. Jedná se o odhad fundamentální frekvence ze spektra SPECF0, cepstra CEPSF0 a autokorelační funkce ACRF0.

3.2 Parametrizace časového vývoje parametru

Časový průběh výše popsaných parametrů je popsán dvěma dalšími parametry: LRE (*Line Regression Error*) a CL (*Curve Length*).

Parametr LRE vyjadřuje rozdíl funkčních hodnot a regresní přímky dané funkce.

$$\text{LRE}(n) = x(n) - (\alpha n + \beta), \quad (1)$$

kde $x(n)$ je časový průběh analyzované funkce a vztah $\alpha n + \beta$ je rovnice popisující průběh regresní přímky, který získáme v prostředí MATLAB např. funkcí POLYFIT. Tímto postupem analyzovaný průběh zbavíme trendu (lineární závislosti) a ze signálu odstraníme stejnosměrnou složku. Funkci LRE je možné dále parametrizovat, např. hledat maximální odchylku od lineární aproximace, spočítat směrodatnou odchylku chyby aproximace (její výkon) apod.

Průměrnou absolutní změnu v signálu popisuje parametr CL [15]

$$CL = \sum_{n=0}^{N-1} |x(n) - x(n+1)|. \quad (2)$$

Ze vztahu je patrné, že signál s více změnami bude dosahovat vyšších hodnot – průměrná hodnota absolutní hodnoty difference bude vyšší.

4 Výsledky

Výsledky jsou prezentovány na parametru GNE (Glottal-to-Noise Excitation Ratio) – byla vy počtena míra korelace mezi průměrným hodnocením odborníků a parametrem GNE. V Obr. 2 je ukázka zvukové nahrávky prodloužené fonace hlásky /a/ pro zdravého pacienta (Obr. 2(a)) a pro pacienta s chraptivostí (Obr. 2(b)) a dále průběh parametru GNE.

Tab. 1 obsahuje souhrnné výsledky Spearmanova korelačního koeficientu ρ pro celou databázi promluv. Z tabulky je patrné, že popis časového vývoje parametru přináší jistou novou informaci a např. pro hodnocení dyšnosti B je výhodnější použít popis délky křivky parametrem CL než prostým průměrem.

Tabulka 1: MÍRA KORELACE PARAMETRU GNE S PRŮMĚRNÝM HODNOCENÍM ODBORNÍKŮ.

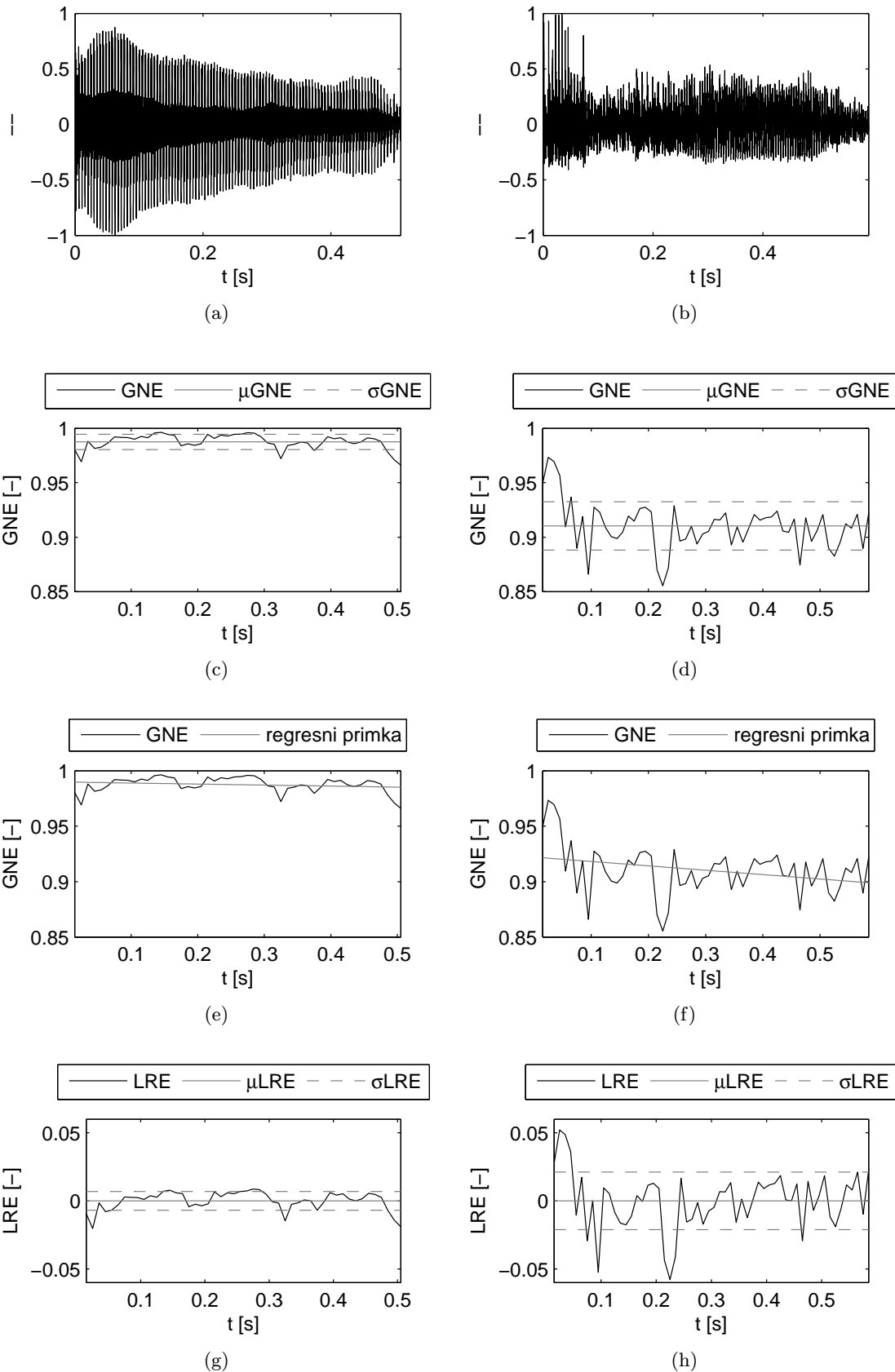
parametr	G	R	B	T
μGNE	-0,71	-0,58	-0,75	0,39
med(GNE)	-0,71	-0,58	-0,74	0,37
σGNE	-0,56	-0,42	-0,67	0,41
$SD(GNE)$	-0,25	-0,13	-0,40	0,28
$CL_{LRE_{GNE}}$	-0,69	-0,58	-0,79	0,36
$\sigma_{LRE_{GNE}}$	-0,58	-0,43	-0,70	0,41
max(LRE _{GNE})	-0,51	-0,38	-0,64	0,38
α_{GNE}	0,14	0,14	0,08	0,01
$SD(LRE_{GNE})$	0,04	0,06	0,01	0,05
$\mu_{LRE_{GNE}}$	0,03	0,01	0,02	0,03
med(LRE _{GNE})	0,01	0,01	0,09	-0,08

5 Závěr

V rámci experimentu bylo provedeno porovnání, zda podrobnější popis parametrizace akustického signálu zlepší korelaci mezi objektivním (automatickým) a subjektivním hodnocením chraptivosti, než při použití pouze střední hodnoty a směrodatné odchylky. K podrobnějšímu popisu časového vývoje parametrů byly použity dva přístupy: hodnocení délky křivky parametrem CL a parametr LRE, který vyjadřuje rozdíl funkčních hodnot a regresní přímky dané funkce. Výsledky jsou prezentovány na průběhu parametru GNE, kde při hodnocení dyšnosti B a při použití parametru CL dojde ke zvýšení míry korelace: $\rho_{B_\mu} = -0,75$, resp. $\rho_{B_{CL}} = -0,79$.

Poděkování

Rádi poděkovali logopedům z Foniatrické kliniky 1. LF UK a VFN v Praze za jejich hodnocení chraptivých promluv, konkrétně Dr. Ing. Janu Vokřálovi, Mgr. Jitce Mercelové, Mgr. Jaroslavě Mrkvičkové, Mgr. Barboře Richterové a Mgr. Janu Dezortovi. Tato práce byla podporována z grantů GACR 102/12/2230 a SGS12/185/OHK4/3T/13.



Obrázek 2: Porovnání průběhů parametru GNE pro zdravého pacienta (a), (c), (e), (g) a pro pacienta s chrapotivostí (b), (d), (f), (h). Na (a) a (b) jsou časové průběhy prodloužené fonace /a/, na (c) až (h) parametrizovaný průběh parametru GNE.

Reference

- [1] James HILLENBRAND and Robert A. HOUDE. Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. *J Speech Hear Res*, 39(2):311–321, 1996.

- [2] Katrin WERTH, Daniel VOIGT, Michael DÖLLINGER, Ulrich EYSHOLDT, and Jörg LOHSCHELLER. Clinical value of acoustic voice measures: a retrospective study. *European Archives of Oto-Rhino-Laryngology*, 267(8):1261–1271, 2009.
- [3] Marc S. De BODT, Floris L. WUYTS, Van de HEYNING Paul H., and Christophe CROUX. Test-retest study of the grbas scale: Influence of experience and professional background on perceptual rating of voice quality. *Journal of Voice*, 11(1):74–80, 1997.
- [4] Michael P. KARNELL, Sarah D. MELTON, Jana M. CHILDES, Todd C. COLEMAN, Scott A. DAILEY, and Henry T. HOFFMAN. Reliability of clinician-based (grbas and cape-v) and patient-based (v-rqol and ipvi) documentation of voice disorders. *Journal of Voice*, 21(5):576–590, 2007. doi: DOI: 10.1016/j.jvoice.2006.05.001.
- [5] Floris L. WUYTS, Marc S. De BODT, and Van de HEYNING Paul H. Is the reliability of a visual analog scale higher than an ordinal scale? an experiment with the grbas scale for the perceptual evaluation of dysphonia. *Journal of Voice*, 13(4):508–517, 1999. doi: DOI: 10.1016/S0892-1997(99)80006-X.
- [6] Kumara SHAMA, Anantha KRISHNA, and Miranjan U. CHOLAYYA. Study of harmonics-to-noise ratio and critical-band energy spectrum of speech as acoustic indicators of laryngeal and voice pathology. *EURASIP J. Appl. Signal Process.*, pages 50–50, 2007.
- [7] Peter J. MURPHY. Perturbation-free measurement of the harmonics-to-noise ratio in voice signals using pitch synchronous harmonic analysis. *The Journal of the Acoustical Society of America*, 105(5):2866–2811, 1999. doi: 10.1121/1.426901.
- [8] Eiji YUMOTO, Yumi SASAKI, and Hiroshi OKAMURA. Harmonics-to-noise ratio and physiological measurement of the degree of hoarseness. *JSHLR*, 27:2–6, 1984.
- [9] B. YEGNANARAYANA, Christophe d’ALESSANDRO, and Vassilis DARSINOS. An iterative algorithm for decomposition of speech signals into periodic and aperiodic components. *IEEE Transactions on Speech and Audio Processing*, 6(1):1–11, 1998.
- [10] Peter J. MURPHY and Olatunji O. AKANDE. Quantification of glottal and voiced speech harmonics-to-noise ratios using cepstral-based estimation. In *NOLISP*, volume 3817, pages 150–160, 2005. doi: http://dx.doi.org/10.1007/11613107_13.
- [11] Claudia MANFREDI. Adaptive noise energy estimation in pathological speech signals. *Biomedical Engineering, IEEE Transactions on*, 47(11):1538–1543, 2000. doi: 10.1109/10.880107.
- [12] D. MICHAELIS, T. GRAMSS, and H. W. STRUBE. Glottal-to-noise excitation ratio – a new measure for describing pathological voices. *ACUSTICA*, 83:700–706, 1997.
- [13] Juan I. GODINO-LLORENTE, Víctor OSMA-RUIZ, SÁENZ-LECHÓN Nicolás, Pedro GOMEZ-VILDA, Manuel BLANCO-VELASCO, and Fernando. CRUZ-ROLDÁN. The effectiveness of the glottal to noise excitation ratio for the screening of voice disorders. *Journal of Voice*, 24(1):47–56, 2010.
- [14] A. FRID and Y. LAVNER. Acoustic-phonetic analysis of fricatives for classification using svm based algorithm. In *Electrical and Electronics Engineers in Israel (IEEEI), 2010 IEEE 26th Convention of*, pages 751–755, 2010. doi: 10.1109/EEEI.2010.5662110.
- [15] Yaozhang Pan, Shuzhi Sam Ge, Feng Ru Tang, and A. Al Mamun. Detection of epileptic spike-wave discharges using svm. In *Control Applications, 2007. CCA 2007. IEEE International Conference on*, pages 467–472, oct. 2007. doi: 10.1109/CCA.2007.4389275.

Adam Stráník
stranada@fel.cvut.cz

Roman Čmejla
cmejla@fel.cvut.cz