

VZDÁLENÁ ANALÝZA LIDSKÉHO HLASU V REÁLNÉM ČASE.

Tomáš Kulhánek^{1,2}, Marek Frič³, Jaroslav Hrb³

Abstrakt

Během foniatrického vyšetření se provádí analýza hlasu, jejímž výsledkem je diagram hlasového pole, tzv. fonetogram. Aplikace RealVoiceLab vyvinutá centrem hudební akustiky HAMU v prostředí Matlab umožňuje vytvoření fonetogramu v reálném čase a další následné analýzy jakými jsou plocha hlasového pole, poloha nejčtenějších zastoupení, statistiky hlasových projevů po dokončení nahrávání. V příspěvku představíme vzdálené nasazení aplikace RealVoiceLab přes upravený protokol RDP. Tato aplikace je přístupná v síti CESNET 2 a obecně Internetu a umožňuje aplikaci používat pomocí standardního klientského programu „Vzdálená plocha“ nebo i „Remote Desktop“, která přeměrovává grafické rozhraní, vstupy z klávesnice a myši a nahrávaný zvuk k výkonnému serveru, kde aplikace ve skutečnosti běží. Aplikaci používají u specialistů na foniatrii, kteří ji používají k analýze hlasu v reálném čase za přítomnosti pacienta a také k analýze záznamu hlasu archivovaných.

Klíčová slova:

fonetogram, vzdálená plocha, RDP, přenos zvuku

Úvod

V rámci spolupráce Hudební a taneční fakulty akademie múzických umění (HAMU), První lékařské fakulty Univerzity Karlovy v Praze a sdružením CESNET z.s.p.o. vzniká produkt pro podporu vzdálených vyšetření v oblasti ORL medicíny, konkrétně v oblasti pořízení a analýza hlasu vydávaný hlasivkami. Ve výzkumném centru hudební akustiky HAMU vzniká aplikace RealVoiceLab, která provádí analýzu hlasového signálu v real-time režimu, tzn. během nahrávání se zobrazuje rychlá analýza hlasového pole.

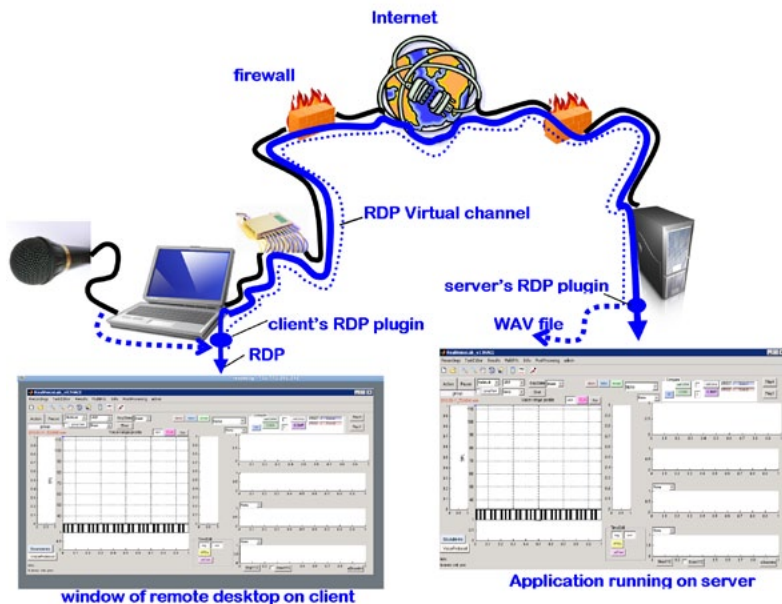
Na konferenci MEDSOFT 2010, jsme představili aplikaci ParVRP[1], která analýzu prováděla v tzv. postprocesingu, tzn. bezprostředně po pořízení nahrávky přes tzv. vzdálenou plochu (Remote Desktop). Vzdálená plocha dovoluje použít aplikaci z jakéhokoliv počítače připojeného na dálku a vybaveného jen klientskou aplikací pro vzdálenou plochu bez nutnosti instalace výpočetních knihoven a konfigurace analytické aplikace v lokálním počítači [2]. Aplikaci jsme rozšířili o možnost sledování a vyhodnocování v reálném čase. Po ukončení nahrávání, lze provádět detailnější analýzy a připravit výsledný report. Aplikace zároveň umožňuje sdílet na souborové úrovni data z lokálního počítače opět přes protokol RDP.

¹LF UK, ²CESNET z.s.p.o., ³HAMU

Metody

Protokol Vzdálené plochy, tzv. RDP přenáší od klientské aplikace (většinou okna) události z klávesnice a změny polohy myši k serveru, která tyto události vyhodnotí a vykreslí je v prostředí, či předá aplikaci. Změny grafického vzhledu jsou zakódovány a přeneseny zpět ke klientovi, který je vykreslí na obrazovku uživatele. Protokol RDP verze 5.2 umí přesměrovat zvuk ke vzdálené aplikaci ke klientovi na zvukovou kartu, ale neumí přesměrovat nahrávání zvuku z klientského mikrofону k serverové aplikaci. Protokol RDP verze 7 dostupný v novějších verzích MS Windows 2008 R2 a Windows 7, případně pluginy třetích stran umí přesměrovat i nahrávání zvuku ke vzdálené aplikaci běžící na serveru. Při přenosu zvuku touto cestou jsou však použity zvukové kodeky, které jsou ztrátové a v předchozím výzkumu se ukázaly jako nevhodné pro přesnou analýzu lidského hlasu. Proto jsme vyvinuli vlastní metodu přenosu nahrávaného zvuku. Plugin v protokolu RDP, který je naprogramován v technologii .NET, využívá samostatné kanály RDP spojení tak, aby po nich posílal a přijímal proprietární data (v našem případě řídicí signály pro spuštění) a ukončení nahrávání a binární audio data ve formátu WAV. Kvalita přenášeného zvuku je tak pod kontrolou analytické aplikace RealVoiceLab[2].

Zvukové vzorky se v nahrávce několikrát opakují, analýza několika vybraných vzorků během nahrávání proto poskytne dostačující přehled o hlasovém rozsahu. Analytická část aplikace v reálném čase vyhodnocuje jen několik



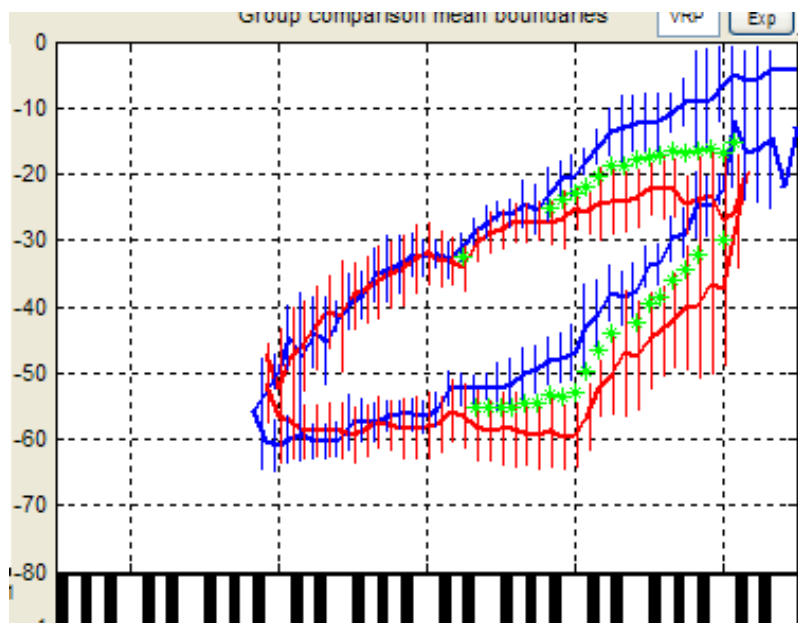
Obrázek 1 — Obrazovka aplikace RealVoiceLab s analýzou hlasového pole

vybraných vzorků zvuku za sekundu podle rychlosti s jakou je stačí analyzovat. Po dokončení nahrávání se analyzuje zbytek nahrávky už post-procesingově. Plugin protokolu RDP využívající technologii .NET byl pro analýzu v reálném čase upraven tak, aby poskytoval vybrané vzorky matematické části aplikace k rychlé analýze z datového toku, který proudí z vyhrazeného RDP kanálu.

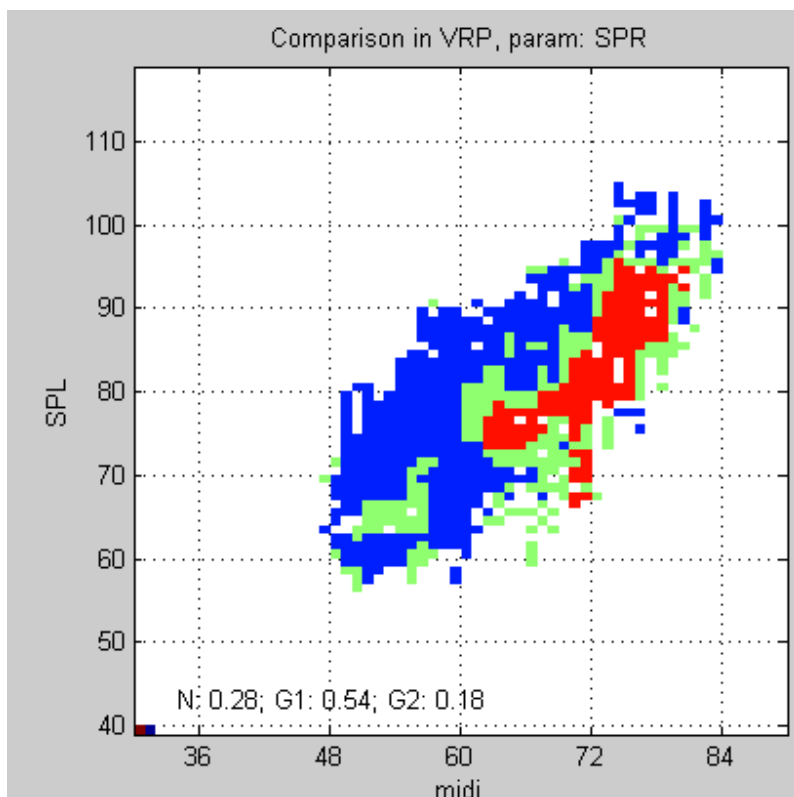
Analýza zvuku

Aplikace RealVoiceLab využívá numerické knihovny systému Matlab a ze zvukových dat ve formátu WAV analyzuje základní frekvenci (pomocí autokorelační funkce), vypočte okamžitou energii signálu a spektrální charakteristiky pomocí rychlé Fourierovy transformace. Vypočtená data jsou následně efektivně filtrována podle vlastností lidského hlasu, tak aby vytvořila konečné zobrazení hlasového pole (voice range profile, VRP) a jeho parametrů.

Na virtuálních serverech se systémy MS Windows 2003 a MS Windows 2008 provozujeme aplikaci RealVoiceLab, která se automaticky spustí při přihlášení uživatele na vzdálenou plochu, připojí lokální disk a nastaví přeměrování nahrávání k lokálnímu mikrofonu.



Obrázek 2 — Porovnání průměrných obrysů zpěvních hlasových polí pěvecky trénovaných (modře) a netrénovaných (červeně) studentek. Zelené hvězdičky označují tóny se statisticky významným rozdílem v průměrně dosažené hladině akustického tlaku.



Obrázek 3 — Porovnání rozložení parametru Singing power ratio s ohledem na pozici v hlasovém poli. Modré části ukazují část překrývajícího se hlasového pole s naměřeným parametrem statisticky významně vyšším pro první skupinu, červeně pro druhou skupinu, zelené části nevykazují významné rozdíly.

Výsledky

Pokud je nahrávka pořízena kalibrovaným mikrofonem v předem dané vzdálenosti od obličeje, pak lze poměrně dobře porovnat analýzy hlasů různých lidí a najít statisticky významné rozdíly, které mohou sloužit k obecným závěrům či dalšímu výzkumu. Přípravek k upevnění mikrofonu na hlavu je doporučován uživatelům programu.

Vzdálený přístup k analýze hlasu pomocí programu RealVoiceLab a serveru medicus.cesnet.cz byl využit v roce 2011 k několika vědeckým účelům. Významným příkladem spolupráce je výzkum Mgr. A. Kulanové z Katolické Univerzity v Ružomberku, která systém použila pro zpracování akustických nahrávek vzorku svých studentů, které vzdělává v oblasti liturgického zpěvu. Na základě uvedené spolupráce vznikla studie, připravená k publikování

v časopise Disputationes (SK), pod názvem: „A. Kulanová, M. Frič: Komparácia náčuvových a počítačových akustických hodnotení speváckych vlastností hlasu“. Dále byl systém použit při výzkumu porovnání vlastností hlasu pěvecky trénovaných a netrénovaných studentů pedagogiky na Pedagogické fakultě Univerzity Karlovy [3], viz též obrázek 2 a 3.

Závěr

Přenos zvuku pořízeného na lokálním počítači ke vzdálené aplikaci pomocí virtuálních kanálů v protokolu RDP je efektivním řešením pro systémy sledující koncept tenkého klienta i pro systémy, které chtějí obecně přenést balík dat z klienta na server. Využití této koncepce pro nasazení aplikace RealVoiceLab vede k jejímu použití z různých míst, kterým stačí připojení k Internetu rychlostí aspoň 1MBit/s.

Poděkování

Práce na tomto projektu byla podpořena z projektu Velká Infrastruktura CESNET a z fondu rozvoje sdružení CESNET č. 423 r. 2011.

Literatura

- [1.] Frič M., *Parametrizovaný fonetogram obecných řečových a hlasových projevů – ParVRP, MARC–Technologický list čís. 12, Zvukové studio HAMU, Praha 2007, <http://web.hamu.cz/zvuk/vyzkum/dokumenty/TL12x.pdf>*
- [2.] Kulhánek T., Frič M., Šárek M.: *Vzdálená analýza lidského hlasu – bezeztrátové nahrávání zvuku přes IP síť, MEDSOFT 2010, sborník příspěvků, vydal Creative Connections s.r.o., ISSN 1803–8115*
- [3.] Králová M.: *Nároky učitelské profese na hlas pedagoga, diplomová práce, Pedagogická fakulta Univerzity Karlovy v Praze, katedra hudební výchovy, Praha 2011*

Kontakt:

Mgr. Tomáš Kulhánek

Ústav patologické fyziologie 1.LFUK

U Nemocnice 5

128 53 Praha 2

email: tomas.kulhanek@lf1.cuni.cz