

3D Elektromagnetický artikulograf

Martin Matura¹

1 Úvod

Elektromagnetický artikulograf (viz obr. 1) umožňuje digitální nahrávání, reprezentaci a vyhodnocení pohybů artikulátorů při vytváření řeči. Mezi artikulátory řadíme nejčastěji špičku a tělo jazyka, rty, dolní čelist a měkké patro. Princip fungování přístroje je založen na indukčním měření vzdáleností, přičemž je důležité, aby hlava subjektu byla umístěna uvnitř sférické měřící oblasti pod vysílacími cívkami, jinak by mohlo dojít ke zkreslení výsledků.

K artikulátorům jsou připevněny senzory v podobě malých cívek pomocí fyziologického lepidla a v horní části artikulografu je umístěno devět vysílacích cívek. Každá z nich produkuje střídavé magnetické pole na různé frekvenci, které indukují střídavý proud v senzorech. To můžeme zaznamenat jako analogové signály. Síla indukovaného proudu je pak funkcí vzdálenosti a úhlu senzoru k vysílací cívce. Proto, a díky různým frekvencím vysílacích cívek, je možné zároveň získat amplitudy každého senzoru ze všech vysílacích cívek. Kombinace všech devíti amplitud (jedna z každé vysílací cívky) nám potom umožňuje vyhodnotit pozici a orientaci každého senzoru.

Samotné snímání pohybu artikulátorů by nám však nestačilo, protože bychom neměli informaci o vyslovené promluvě a nemohli bychom tak k jednotlivým zvukům přiřadit pohyby provedené při jejich vytváření. Proto dochází i k nahrávání akustického signálu a jeho následné synchronizaci v externím zařízení s daty z artikulografu. Důležitou vlastností 3D artikulografu je také schopnost odstranit nežádoucí pohyby hlavy při mluvení, které by jinak znehodnotily výsledné trajektorie artikulátorů.



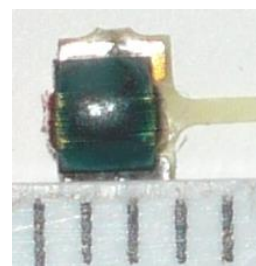
Obrázek 1: Elektromagnetický artikulograf

2 Senzory

Jak už bylo řečeno, senzory používáme jednak pro snímání pohybu artikulátorů a dále pak máme dva až tři senzory pro snímání pohybů hlavy, tzv. referenční senzory.

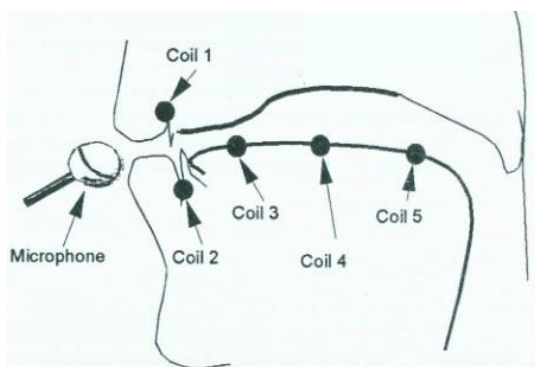
¹ student doktorského studijního programu Aplikované vědy a informatika, obor Kybernetika, e-mail: mate221@kky.zcu.cz

Referenční senzory jsou připevněny na místech, které se při mluvení nepohybují (horní čelist, nosní můstek nebo za ušima) a k jejich pohybu tudíž dochází pouze při pohybu hlavy. Největším úskalím při použití sensorů, které mají malé rozměry 2x3 mm (obr. 2), aby při mluvení nepřekážely, je jejich připevnění. Snahou je, aby senzory vydržely přilepeny co nejdéle, abychom mohli nahrát dostatečné množství dat. Už při prvních pokusech však bylo zřejmé, že problematické bude upevnění sensorů na jazyku.



Obrázek 2: Senzor

Při běžném nahrávání by měl mluvčí během nahrávacího procesu, který trvá většinou několik hodin, namluvit minimálně 500 vět a v ideálním případě pak kolem 1000 vět. U našeho prvního testování sensorů jsme zkoušeli pouze tři, které jsme přilepili postupně na kořen, tělo a špičku jazyka. První senzor bohužel odpadl již po 20 minutách, druhý po 45 minutách a třetí zhruba po hodině. Z pokusů, které provedl Richmond et al. (2011) vyplynulo, že byli schopni nahrávat několik hodin, proto nás naše časy velmi překvapily.



Obrázek 3: Umístění sensorů

příprava plochy jazyka je důležitá a velmi ovlivňuje maximální možnou dobu nahrávání.

U dalšího testování jsme použili pět sensorů (obr. 3) a kladli jsme větší důraz na přípravu plochy jazyka, abychom zajistili, že lepidlo dobře přilne a udrží senzory po delší dobu. Nejprve jsme očistili vrchní vrstvu jazyka zubním kartáčkem, poté jsme ho vysušili lékařskou gázou, očistili alkoholem, znovu vysušili a nakonec přilepili senzory. Celý proces připevnění sensorů trval asi patnáct minut. V průběhu nahrávání odpadl jako první, kvůli přílišnému tření drátků a kontaktu se zuby, senzor na špičce jazyka, který vydržel jen 30 minut. Zbylé 2 senzory vydržely přilepeny téměř tři hodiny, což ukázalo, že důkladná

3 Plánované experimenty

Výstupem artikulo grafu jsou artikulační trajektorie společně s informací o zvuku, kterému patří. Toho bychom rádi využili v konkatenační syntéze řeči, kterou provádíme pomocí metody výběru jednotek. U této metody dochází k vybírání segmentů řeči z velké řečové databáze a k jejich následnému řetězení. Artikulační trajektorie by se tak mohly stát jedním z elementů, které by přispěly k výběru lepších řečových segmentů, což testovali Richmond a King (2016), čímž by se zvýšila kvalita syntetické řeči.

Poděkování

Příspěvek byl podpořen grantovým projektem SGS-2016-039.

Literatura

- Richmond, K., Hoole, P., a King, S., 2011. Announcing the electromagnetic articulatory graphy (day 1) subset of the mngu0 articulatory corpus. *Proceedings Interspeech*, pp 1505-1508.
- Richmond, K., a King, S., 2016. Smooth Talking: Articulatory Joint Cost for Unit Selection. *Proceedings, 41st Int. Conference on Acoustics, Speech and Signal Processing*.