

REFERENCE SIGNALS IN INTRACRANIAL EEG: IMPLEMENTATION AND ANALYSIS

Daniel Uher

Bachelor Degree Programme (3), FEEC BUT

E-mail: xpuherd02@vutbr.cz

Supervised by: Marina Ronzhina

E-mail: ronzhina@feec.vutbr.cz

Abstract: The idea of an artifact-free brain activity recording has been circling around the scientific world for a few decades. Noise present in brain activity recordings may complicate the process of evaluation and interpretation. For the elimination of such unwanted components, the concept of virtual reference signals is usually used. In this work, the algorithms for reference signal estimation using common average-based method as well as more recent methods based on independent component analysis (ICA) were realized and evaluated on a new set of real clinical data. It was found that the ICA-based algorithms allow obtaining more accurate estimation of the reference signal as compared to the average-based one. Finally, all the methods were implemented into a free installable Python toolbox, which will be publicly available after additional testing on real data.

Keywords: Reference signals, reference, intracranial EEG, ICA, average, correlation

1 ÚVOD

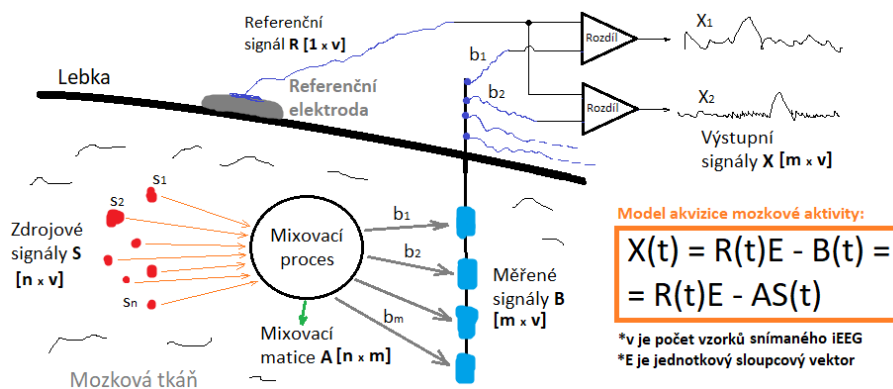
Intrakraniální elektroencefalogram (iEEG) je invazivní obdoba skalpového EEG, při které se měří elektrické potenciály v mozku při přímém kontaktu mezi elektrodou a mozkovou tkání. Parazitní jevy a nežádoucí složky jsou neodmyslitelnou součástí záznamu iEEG a s tím také přicházejí na scénu metody pro jejich co nejúčinnější potlačení a naopak zvýraznění užitečné informace signálu. Jedním z populárních způsobů je identifikace tzv. referenčního signálu. Referenční signál (zkr. reference) je vektor hodnot elektrického potenciálu vyskytujícího se na referenční elektrodě. Tento potenciál není možné měřit přímo, musí být tedy posléze dopočítán (tzv. virtuální referenční signál). V oblasti iEEG se pro tento účel ověřilo použití metod založených na analýze nezávislých komponent (ICA), které sice vykazují nejlepší výsledky, ale jsou výpočetně náročnější a často jsou nahrazeny pouhou "průměrnou referencí". Je třeba také zdůraznit, že matematickými výpočty získáme pouze odhady reálných hodnot a jako takové budou vždy obsahovat určitou odchylku od skutečnosti.[1]

2 METODY ODHADU VIRTUÁLNÍHO REFERENČNÍHO SIGNÁLU

Model vzniku potenciálů snímaných intrakraniální metodou je zobrazen na obrázku č. 1. Každý z měřených potenciálů b je tvořen lineární kombinací zdrojových mozkových signálů s . Tato kombinace je matematicky vyjádřena tzv. mixovací maticí A . Na základě úvahy, že každému aktivnímu kontaktu na elektrodě bude přiřazen jeden jediný dominantní zdroj a tím pádem bude platit $B = S$, můžeme přepsat modelovou rovnici z obrázku č. 1 následujícím způsobem [1]:

$$X = \begin{bmatrix} 1 \\ : \\ 1 \end{bmatrix} - A \quad Z = QZ \quad \text{kde} \quad Z = \begin{bmatrix} R \\ S \end{bmatrix} \quad (1)$$

kde Q je mixovací matice a Z je matice skládající se ze submatice zdrojových signálů S a referenčního vektoru R .



Obrázek 1: Model vzniku a akvizice elektrických potenciálů mozku

Analýza nezávislých komponent je vícerozměrná statistická metoda, která dokáže identifikovat zdrojové signály z měřeného setu výstupních signálů X . Metoda podrobně popsána v [1]. Spočívá v nalezení tzv. demixovací matice W a aplikována na matici X jako $\bar{Z} = WX$, kde \bar{Z} je výsledná matice nezávislých komponent (zkr. ICs), která obsahuje hledaný odhad referenčního signálu R . Pro účely této práce je důležité poznamenat, že ICA má zpravidla jedno jediné řešení, pouze pokud je počet detektorů (v tomto případě aktivních kontaktů elektrody) větší nebo stejný jako počet zdrojů (zdrojových signálů). V opačném případě bude existovat více možných řešení.[1][2] Pro účely této práce byla použita metoda FastICA z balíčku *Scikit-learn* [2] s parametry ponechanými na defaultních hodnotách (maximální počet iterací 1000 a použití paralelních výpočtů), což se ukázalo jako nejefektivnější. Součástí je také předzpracování v podobě bělení (dekorelace) vstupních dat X . Spolehlivost odhadu nezávislých komponent zde testována nebyla.

2.1 METODA PRŮMĚROVÁNÍ

Nejběžněji používanou metodou pro získání referenčního signálu je tzv. společný průměr (angl. common average reference). Princip spočívá ve vytvoření průměrného signálu ze setu měřených signálů iEEG. Tím se zvýrazní společné složky, zatímco ty jedinečné budou potlačeny. V této práci je průměrná reference R_{avg} použita pro posouzení kvality dopočítaných referencí z ostatních metod.

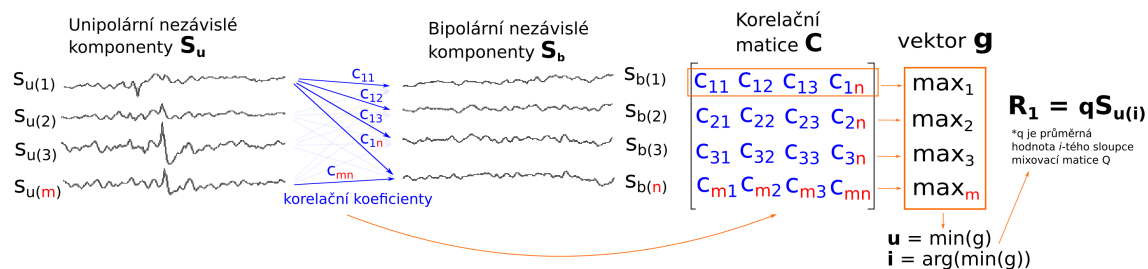
2.2 METODA ZALOŽENÁ NA ICA A KORELAČNÍ ANALÝZE

Na obr. č. 2 je znázorněn princip odhadu reference za použití ICs spočtených pro unipolární a bipolární data. Vychází se z předpokladu, že referenční signál se bude vyskytovat pouze v setu ICs z unipolární montáže, zatímco v ICs z montáže bipolární se vyskytovat nebude. Tudíž referenční signál je odvozený z komponenty z S_u , která je nejméně korelována s komponentami z S_b . Jelikož metoda má více možných řešení (viz poznámka nahoře), proces hledání R_1 se iteračně opakuje, dokud nalezené minimum u nenabude hodnoty menší než uživatelsky nastavené kritérium p (podle [1] pro kvalitní odhad R_1 platí $u < p$, kde $p = 0.25$).[1]

2.3 METODA ZALOŽENÁ NA ICA A VLASTNOSTECH BIPOLÁRNÍCH DAT

Oproti předchozí metodě se v této aplikuje ICA pouze na bipolární data, u nichž je splněna podmínka, že počet detektorů se rovná počtu zdrojů (viz poznámka nahoře). Akvizice referenčního signálu zde spočívá pouze v aplikaci následující rovnice na výstupní signály X [1]:

$$R_2 = \frac{1}{n} \sum_{i=1}^n R_i = \frac{1}{n} \sum_{i=1}^n \left[x_i - \sum_{l=1}^k \frac{E[x_i s_{b(l)}]}{E[s_{b(l)}^2]} s_{b(l)} \right] \quad (2)$$



Obrázek 2: Princip metody založené na ICA a korelační analýze

kde k je počet bipolárních komponent a E značí průměrnou hodnotu. Odvození vzorce vychází z podmínky, že průměr ze všech korelačních koeficientů mezi referenčním signálem a každou s_b je roven nule. [1]

3 VÝSLEDKY TESTOVÁNÍ METOD

Metody byly testovány na RAW¹ sEEG 16-kanálových datech o velikosti 500 000 vzorků se vzorkovací frekvencí $f_s = 5000$ Hz (100 s záznamu).

Pro referenční signál R_1 byl získán $u = 0.1682$, což ukazuje na dobrý odhad reference. Na druhou stranu korelační koeficient mezi R_1 a R_{avg} udává relativně nízkou hodnotu 0.2549. Korelační koeficient mezi R_2 a R_{avg} je vyšší a ukazuje 0.4510. Důležitým kritériem pro hodnocení kvality referenčního signálu jsou korelační koeficienty s jednotlivými bipolárními komponentami, které by měly být co nejmenší. Pro R_1 byl maximální korelační koeficient s S_b roven 0.3612, pro R_{avg} dokonce až 0.4286. U R_2 však byla nejvyšší hodnota 4.6336×10^{-9} , což je jednoznačně nejlepší výsledek. Korelační koeficient mezi R_1 a R_2 konverguje k 0.9426, což ukazuje na vysokou podobnost mezi zjištěnými referencemi. Rozdíly v korelačních koeficientech mohou být způsobeny přítomností výrazného trendu v R_1 a R_2 , kdežto u R_{avg} se tento trend mohl vyskytovat pouze částečně.

4 ZÁVĚR

V této práci byly testovány tři metody pro kalkulaci referenčního signálu z měřeného setu iEEG dat. Díky testovacím výsledkům se podařilo prokázat výsledky skupiny Worrell a spol.[1], tedy to, že jako nejspolehlivější odhad skutečného referenčního signálu se jeví R_2 . Toto tvrzení se prokázalo nejenom vyšší korelací s R_{avg} , ale hlavně téměř nulovými korelacemi mezi R_2 a všemi bipolárními ICs. Všechny použité algoritmy byly implementovány do knihovny funkcí s názvem *Refsig*, vytvořené v jazyku Python. Po ověření správné funkčnosti na více testovacích datech bude tato knihovna volně stažitelná a instalovatelná na jakýkoliv počítač s nainstalovaným Pythonem verze vyšší než 2 a přístupem na internet.

REFERENCE

- [1] HU,S., STEAD,M., WORRELL,G.A. *Automatic identification and removal of scalp reference signal for intracranial eegs based on independent component analysis*. IEEE Transactions on Biomedical Engineering, 2007: č. 54, s. 1560–1572.
- [2] PEDREGOSA, F., VAROQUAUX, G., GRAMFORT, A., MICHEL, V. 2011. *Scikit-learn: Machine Learning in Python*. Journal of Machine Learning Research, 12, 2825-2830.

¹Unipolární data, která nebyla nijak dále zpracována