

Seminární práce č. 1

Analýza EKG signálů a korelace měřených signálů

Adam Legner

F22130

Počítačové zpracování signálů (PZS)

Univerzita J. E. Purkyně

Zadání

Seminární práce se skládá ze dvou úkolů.

Úkol 1: Ve zdrojové databázi najdete celkem 17 měření EKG signálu. Signál je již filtrován a centralizován kolem podélné osy. EKG signál obsahuje dominantní peaky, které se nazývají R vrcholy. Vzdálenost těchto vrcholů určuje dobu mezi jednotlivými tepy. Počet tepů za minutu je tedy počet R vrcholů v signálu o délce jedné minuty. Navrhněte algoritmus, který bude automaticky detekovat počet R vrcholů v EKG signálech a prezentujte tepovou frekvenci při jednotlivých jízdách/měřeních. Váš algoritmus následně otestujte na databázi MIT-BIH a prezentujte jeho úspěšnost vzhledem k anotovaným datům z databáze.

Úkol 2: Ve zdrojové databázi najdete celkem 13 pacientů, kde se u každého měří celkem 3 signály (EKG, ABP a ICP). U každého z pacientů vypočítejte korelaci mezi jejich EKG, ABP a ICP signálem. Dále u všech pacientů provedte korelace mezi jejich ABP signály. Před výpočty ověřte, že všechny signály mají stejnou vzorkovací frekvenci a jsou centralizovány podle podélné osy. V opačném případě převzorkujte všechny signály na Vámi zvolenou vzorkovací frekvenci. Pro smysluplné provedení korelační analýzy na ABP signálech je nutné nejprve srovnat signály na stejný počátek, nejlépe dle dominantního vrcholu. Vyhodnoťte vždy pouze první hodinu signálu. Pokud v signálech najdete chybějící data, provedte jejich nahrazení například pomocí interpolace.

Postup

Úkol 1 - Zpracování EKG signálů

Načtení dat

Data jsem načetl pomocí knihovny WFDB, která umí pracovat s formátem PhysioNet. Z DriveDB jsem použil všech 18 záznamů (drive01 až drive17b) – většina má vzorkovací frekvenci 496 Hz, některé 480 nebo 356 Hz. EKG signál je vždy v kanálu 0. Z databáze CHARIS jsem načetl 3 záznamy (charis1-charis3), kde EKG kanál je „ECG“ a vzorkovací frekvence je 50 Hz. Zpracovával jsem první 2 minuty záznamu.

Předzpracování signálu

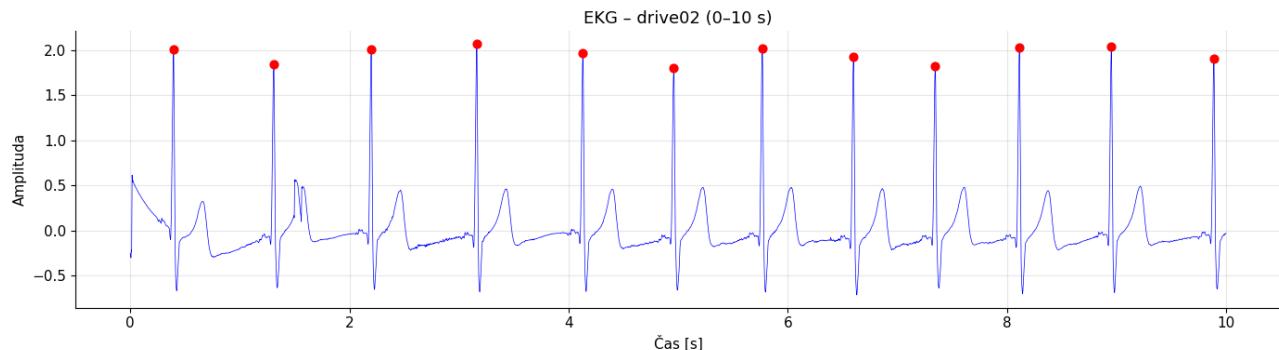
Surový EKG signál obsahuje pomalý drift a stejnosměrnou složku, které je potřeba odstranit.

- 1) Centrování – odečtení průměru signálu, aby se pohyboval kolem nuly.
- 2) RC horní propust – filtr propouštějící frekvence nad 0,5 Hz. Odstraní pomalý drift.
- 3) FFT odstranění driftu – signál jsem převedl do frekvenční oblasti pomocí FFT, vynuloval složky pod 0,5 Hz a zpětným FFT jsem získal vyčištěný signál. Dále jsem kontroloval polaritu – pokud jel signál do záporných hodnot, invertoval jsem ho, aby R-píky směřovaly nahoru.

Detekce R-píků

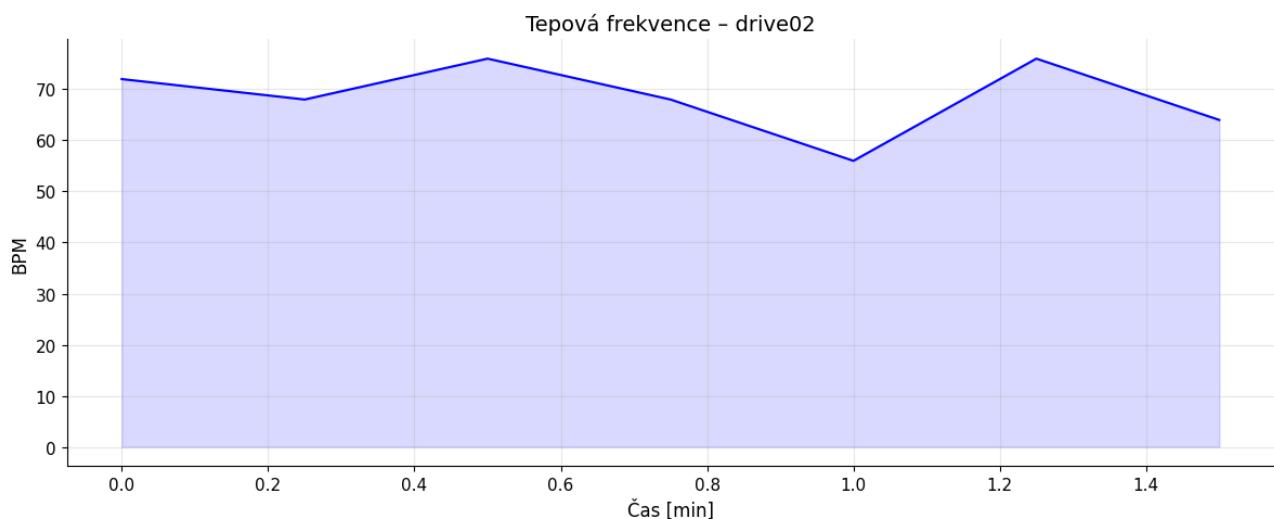
Pro detekci R-píků jsem signál umocnil na druhou, čímž se zvýrazní QRS komplexy a potlačí šum. Jako práh jsem dal 97. percentil umocněného signálu - hodnoty nad prahem jsou kandidáti na R-pík. Z kandidátů jsem vybral pouze lokální maxima se vzájemným odstupem alespoň 0,3 sekundy, kvůli duplicitní detekci.

Na obrázku je prvních 10 sekund záznamu drive02 s označenými R-píky:



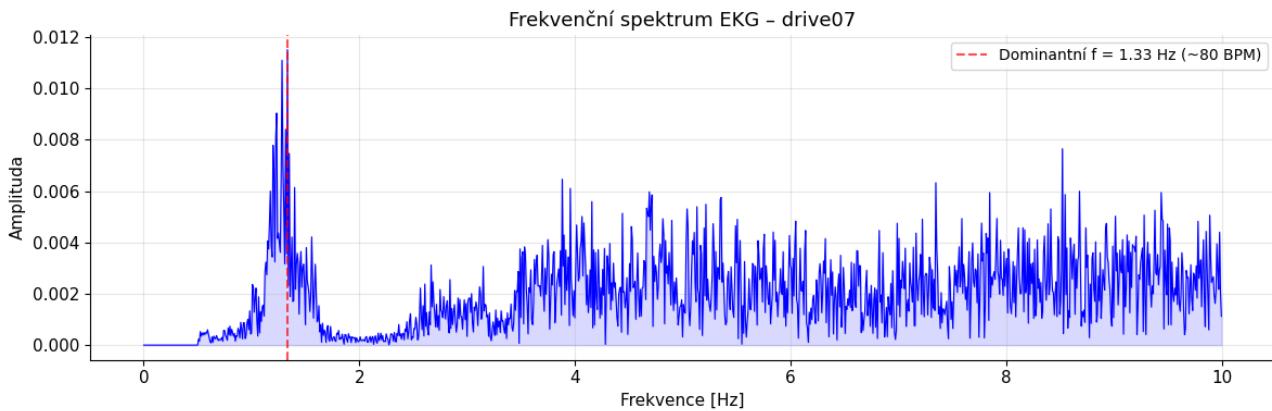
Výpočet tepové frekvence

Ze známých pozic R-píků jsem spočítal RR intervaly (čas mezi po sobě jdoucími R-píky) a z nich vypočítal tepovou frekvenci. Tepová frekvence se vypočítá jako $60 / \text{RR interval}$ a výsledek je BPM. Průběh BPM v čase jsem vykreslil jako graf. U DriveDB záznamů se hodnoty pohybovaly v rozmezí 59–89 BPM, u CHARIS kolem 86–90 BPM.



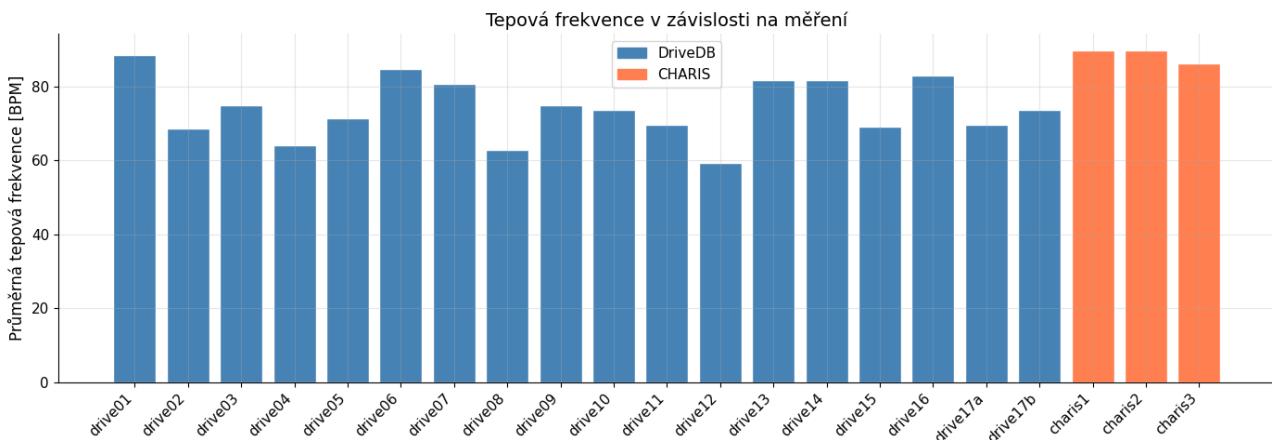
Frekvenční spektrum

Pro každý záznam jsem vykreslil frekvenční spektrum EKG v rozsahu 0–10 Hz pomocí rychlé Fourierovy transformace (FFT) z knihovny `scipy`. Ve spektru jsem nalezl a vyznačil dominantní peak odpovídající tepové frekvenci.



Souhrnné výsledky

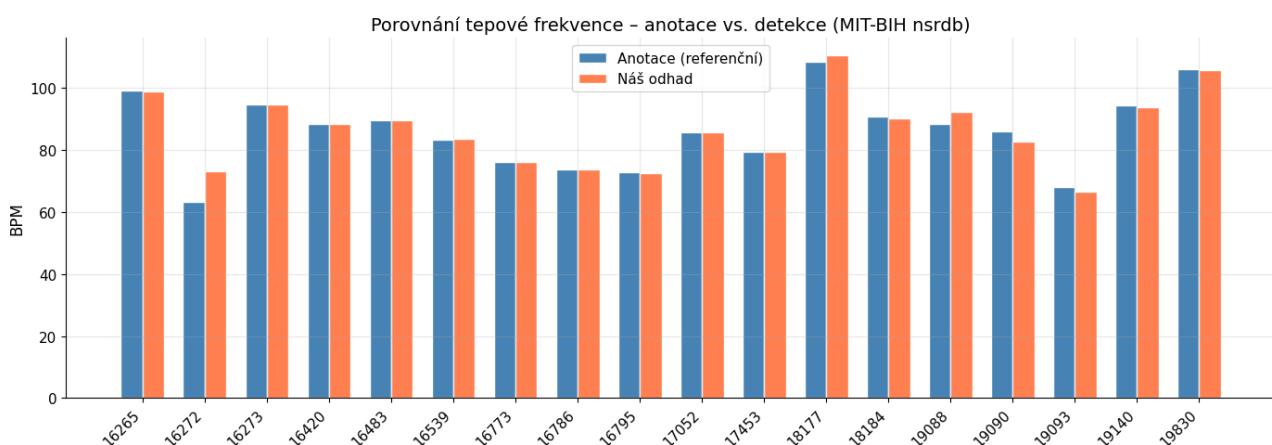
Výsledky všech 21 záznamů (18 DriveDB + 3 CHARIS) jsem shrnul do tabulky s počtem detekovaných R-píků a průměrnou tepovou frekvencí.



Validace na MIT-BIH nsrdb

Správnost detektoru jsem ověřil na databázi MIT-BIH nsrdb, která obsahuje 18 záznamů s referenčními anotacemi R-píků. Detekovaný R-pík jsem bral za správný, pokud ležel do 140 ms od nejbližší referenční značky.

Celková úspěšnost detekce dosáhla 99,77 %. U 7 záznamů z 18 byla úspěšnost 100 %. Nejnižší hodnota 98,76 % (záznam 16265) Velmi dobrý výsledek. Hodnoty BPM spočítané z detekovaných píků odpovídají referenčním anotacím. Na grafu je porovnání tepové frekvence z referenčních anotací a z našeho detektoru pro všechny 18 záznamů:



Úkol 2 - Korelace měřených signálů

Načtení dat z CHARIS

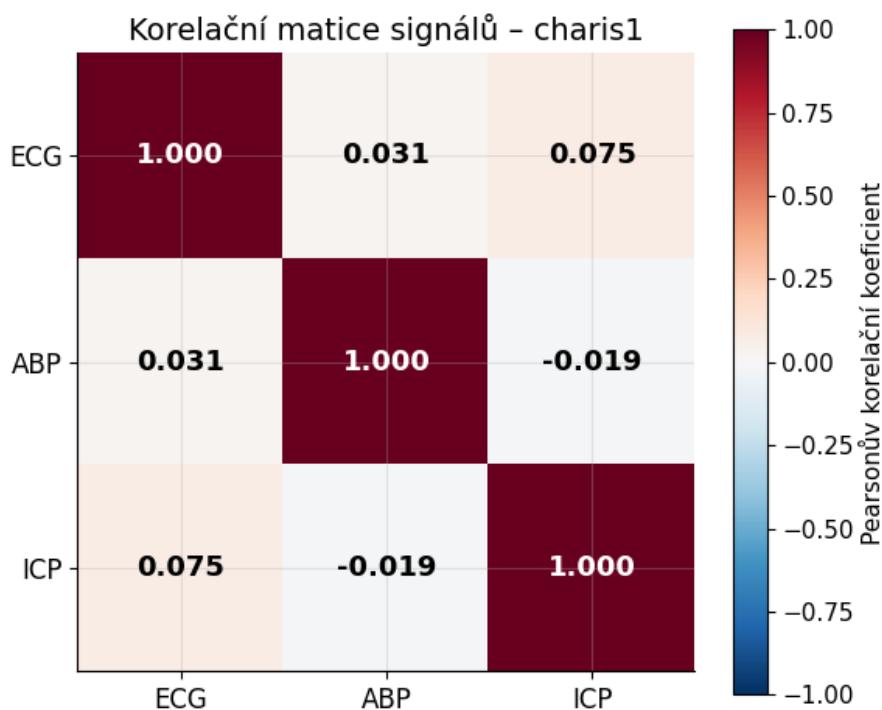
Z databáze CHARIS jsem načetl všech 13 pacientů (charis1 až charis13). Každý záznam obsahuje 3 současně měřené signály: ECG (elektrokardiogram), ABP (arteriální krevní tlak) a ICP (intrakraniální tlak). Všechny signály mají vzorkovací frekvenci 50 Hz. Zpracovával jsem první hodinu každého záznamu, každý signál jsem vycentroval odečtením průměru.

Korelace mezi signály u jednotlivých pacientů

U každého pacienta jsem spočítal Pearsonovu korelacii mezi všemi třemi dvojicemi signálů: ECG-ABP, ECG-ICP a ABP-ICP. Hodnota korelace 1 - úplná shoda, 0 žádná lineární souvislost a -1 opačný průběh.

ECG s tlakovými signály (ABP, ICP) má nízkou korelacii (typicky pod 0,1), což odpovídá tomu, že ECG je elektrický signál a ABP/ICP jsou tlakové veličiny. ABP s ICP korelují hodně - u pacienta charis3 je korelace 0,85, u charis7 hodnoty 0,62.

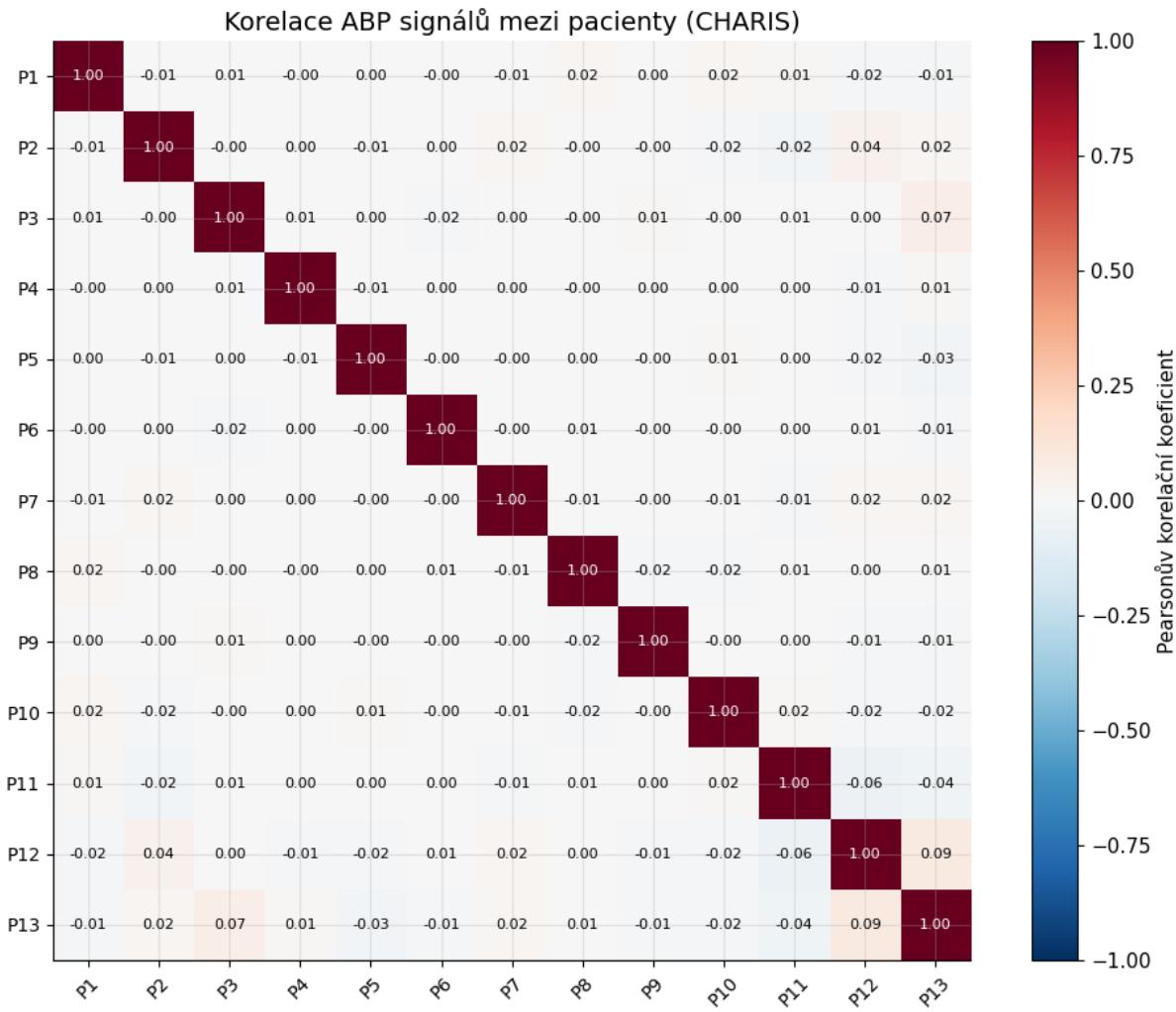
Heatmapa korelací pro pacienta charis1:



Korelace ABP signálů mezi pacienty

ABP signály jsem porovnal mezi všemi 13 pacienty navzájem. Signály jsem nejprve zarovnal - pomocí autokorelace jsem u každého ABP určil délku jedné srdeční periody a signály posunul do shodné fáze.

Finální korel. matice 13×13 ukazuje, že ABP signály pacientů spolu prakticky nekorelují - hodnoty se hýbou od -0,06 do +0,09. To odpovídá individuálním rozdílům v kardiovaskulárním systému - různé parametry jsou u každého pacienta odlišné.



Závěr

Zpracoval jsem EKG signály z databází DriveDB (18 záznamů) a CHARIS (3+13 záznamů). Implementoval jsem detektor R-peaks založený na umocnění signálu, zasazení prahu na 97. percentilu a minimálním odstupu 0,3 s. Předzpracoval jsem data centrováním, RC horní propustí (0,5 Hz) a FFT odstranění driftu.

Detektor jsem validoval na databázi MIT-BIH nsrdb (18 záznamů s referenčními anotacemi) a dosáhl celkové úspěšnosti 99,77 %, přičemž u 7 záznamů byla úspěšnost 100 %. Tepové frekvence z DriveDB se pohybovaly v rozmezí 59–89 BPM, u CHARIS 86–90 BPM.

V druhém úkolu jsem analyzoval korelací tří signálů (ECG, ABP, ICP) u 13 pacientů z databáze CHARIS. Korelace ECG s tlakovými signály byla nízká (pod 0,1), zatímco ABP-ICP dosáhla u některých pacientů hodnotu přes 0,8. Korelace ABP mezi různými pacienty se blíží nule, což sedí k individuálním rozdílům v kardiovaskulárním systému.