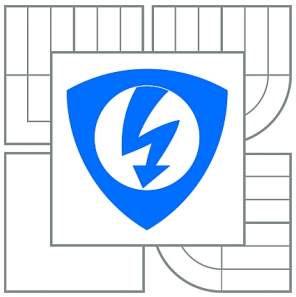
VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

ÚSTAV BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

DEPARTMENT OF BIOMEDICAL ENGINEERING

POUŽITÍ KUMULANTŮ VYŠŠÍCH ŘÁDŮ PRO ANALÝZU EKG

USE OF HIGHER-ORDER CUMULANTS FOR ECG ANALYSIS

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

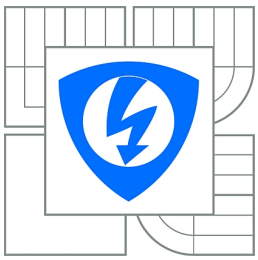
AUTOR PRÁCE — LUCIE MARŠÁNOVÁ

AUTHOR

VEDOUCÍ PRÁCE — Ing. MARINA RONZHINA

SUPERVISOR

BRNO 2013



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky a komunikačních technologií

Ústav biomedicínského inženýrství

Bakalářská práce

bakalářský studijní obor
Biomedicínská technika a bioinformatika

Studentka:	Lucie Maršánová	***ID:***	136481
Ročník:	3	***Akademický rok:***	2012/2013

NÁZEV TÉMATU:

Použití kumulantů vyšších řádů pro analýzu EKG

POKYNY PRO VYPRACOVÁNÍ:

1) Seznamte se s vlastnostmi a způsoby výpočtu kumulantů druhého až čtvrtého řádu. Zaměřte se na využití daných kumulantů pro analýzu EKG. Proveďte literární rešerši v této oblasti. 2) Navrhněte databázi normálních a patologických srdečních cyklů EKG z veřejných databází nebo signálů dostupných na ÚBMI. 3) Realizujte vypočet kumulantů vyšších řádů pro srdeční cykly z navržené databáze. Využijte programové prostředí MATLAB. 4) Proveďte hodnocení vlastností kumulantů vypočtených z EKG (v porovnání s vlastnostmi původních EKG). 5) Proveďte diskusi získaných výsledků. Mimo jiné diskutujte možnosti použití vypočtených kumulantů pro klasifikaci srdečních cyklů.

DOPORUČENÁ LITERATURA:

[1] MENDEL, Jerry M. Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications. In: Proceedings of the IEEE [online]. [cit. 24.1.2013]. 1991, 278-280. Dostupné z: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=75086.
[2] KARIMIFARD, Saeed a AHMADIAN, Alireza. A robust method for diagnosis of morphological arrhythmias based on Hermitian model of higher-order statistics. In: Biomedical engineering online [online]. [cit. 2012-09-04]. Dostupné z: http://www.biomedical-engineering-online.com/content/10/1/22.

Termín zadání:	11.2.2013	***Termín odevzdání:***	31.5.2013

Vedoucí práce: Ing. Marina Ronzhina
Konzultanti bakalářské práce:

prof. Ing. Ivo Provazník, Ph.D.
Předseda oborové rady

UPOZORNĚNÍ:

Autor bakalářské práce nesmí při vytváření bakalářské práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

ABSTRAKT

Tato práce se zabývá využitím kumulantů vyšších řádu pro analýzu EKG. V první části práce je popsán princip elektrokardiografie. Následuje matematické odvození kumulantů, popis jejich vlastností a jejich využití v současné praxi. V další části práce je popsána realizace výpočtu kumulantů z EKG cyklů v programovacím prostředí Matlab. V praktické části práce jsou pomocí simulovaných situací ověřeny předpokládané vlastnosti kumulantů, mezi které patří minimalizace rozdílu amplitudy, časového posunu a potlačení Gaussovského šumu. Tyto vlastnosti umožňují snížení variability cyklů v rámci jedné skupiny a z toho vyplývající snadnější klasifikaci EKG. Následuje výpočet kumulantů reálných EKG cyklů různých skupin. Nakonec jsou pomocí neuronové sítě provedeny dva způsoby klasifikace. Na základě porovnání klasifikační úspěšnosti původních cyklů a jejich kumulantů je zhodnocena využitelnost kumulantů konkrétních řádu pro zpracování reálných EKG signálů.

KLÍČOVÁ SLOVA

Elektrokardiografie, EKG signál, kumulanty vyšších řádů, klasifikace EKG cyklů, umělá neuronová síť

ABSTRACT

This work deals with using higher order cumulants for analysis ECG. In the first part of work is described principle of electrocardiography, followed by matemathical derivation of higher order cumulants, description of their properties and their use in current practice. In the next part of paper is described caltulation higher order statistic of ECG beats in Matlab programming environment. In the practical part of work are tested predicted properties. Distinctive properties are minimalization of amplitude and time shift and Gaussian noise. This properties of higher order cumulants enable lesser variance of beats in one class and easier clasification ECG. Calculation of cumulants from real ECG beats of various groups is then realized. Classification based on the original ECG cycles and cumulants is performed using artificial neural network. Results of these classification approaches are then compared and discussed.

KEYWORDS

Electrocardiography, ECG signal, higher order cumulants, clasification of ECG beats, artificial neutral network

MARŠÁNOVÁ, L. Použití kumulantů vyšších řádů pro analýzu EKG. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2013. 51 s. Vedoucí bakalářské práce Ing. Marina Ronzhina.

Prohlášení

Prohlašuji, že svoji bakalářskou práci na téma Využití kumulantů vyšších řádů pro analýzu EKG jsem vypracovala samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedeného bakalářské práce dále prohlašuji, že v souvislosti s vytvořením éto práce jsem neporušila autorská práva třetích osob, zejména jsem nezasáhla nedovoleným způsobem do cizích autorských práv osobnostních a jsem si plně vědoma následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009Sb.

V Brně dne 29. května 2013 .....

podpis autora

Poděkování

Děkuji vedoucí mé bakalářské práce Ing. Marině Ronzhině za odborné vedení, cenné rady, připomínky a inspiraci při zpracování mé bakalářské práce. Dále děkuji rodičům, kteří mě podporovali po celou dobu mého studia.

Experimentální část této bakalářské práce byla podpořena v rámci projektu Evropského fondu pro regionální rozvoj FNUSA-ICRC CZ.1.05/1.1.00/02.0123.

V Brně dne 29. května 2013 .....

podpis autora

OBSAH

Seznam obrázků **8**

Seznam tabulek **9**

Úvod **10**

1 Elektrokardiografie **11**

1.1 Šíření elektrického vzruchu srdcem 11

1.2 Způsoby záznamu EKG 12

1.2.1 Standardní 12svodové EKG 12

1.2.2 Ortogonální svodový systém 13

1.3 Popis EKG křivky 13

2 Úvod do teorie kumulantů vyšších řádů **15**

2.1 Odvození kumulantů 15

2.2 Tvorba kumulantů ze signálu v praxi 17

2.3 Základní vlastnosti kumulantů 17

3 Využití kumulantů v současné praxi **19**

4 Realizace výpočtu kumulantů vyšších řádů v prostředí Matlab **21**

4.1 Předzpracování elektrokardiogramu 21

4.2 Popis použitých funkcí 25

5 Ověření vlastností kumulantů **30**

5.1 Vliv posunu po ose x (časová osa) 30

5.2 Vliv posunu po ose y (výchylka elektrogramu) 31

5.3 Vliv přítomnosti Gaussovského šumu 31

5.4 Výpočet kumulantů reálných dat 33

6 Statistické zhodnocení variability dat **36**

6.1 Korelační analýza 36

6.1.1 Výsledky korelační analýzy 36

6.2 Příznaky odvozené z kumulantů 37

7 Klasifikace dat pomocí umělé neuronové sítě **41**

7.1 Neuronová síť 41

7.1.1 Obecný popis neuronových sítí 41
7.1.2 Použitá neuronová síť 42
7.2 Klasifikace na základě celých průběhů 43
7.3 Klasifikace na základě příznaků odvozených z kumulantů 44
Závěr 46
Seznam použité literatury 48
Seznam zkratek 51

SEZNAM OBRÁZKŮ

Obr. 1 Šíření elektrického vzruchu v srdci a vznik křivky EKG (převzato a upraveno z [14]) 12

Obr. 2 Rozmístění a zapojení elektrod Frankova ortogonálního systému (převzato a upraveno z [23]) 13

Obr. 3 Fyziologická křivka EKG (převzato a upraveno z [22]) 14

Obr. 4 Cykly z ischemických fázích roztříděné dle morfologie 22

Obr. 5 Cykly z fází reperfuze roztříděné dle morfologie 23

Obr. 6 Cykly z kontrolní fáze roztříděné dle morfologie 24

Obr. 7 Normální cykly ze všech fází roztříděné a sloučené na základě morfologie 25

Obr. 8 Blokové schéma funkce *cum2est* pro výpočet kumulantu druhého řádu 27

Obr. 9 Potlačení posunu po ose x (časový posun) 30

Obr. 10 Potlačení posunu po ose y (výchylka elektrogramu) 31

Obr. 11 Potlačení šumu v signálu 32

Obr. 12 Originální signál 50ti fyziologických EG cyklů a jejich kumulanty 34

Obr. 13 Originální signál 50ti ischemických EG cyklů a jejich kumulanty 34

Obr. 14 Originální signál 61 EG ektopických cyklů s vysokou variabilitou a jejich kumulanty(Jiný typ) 35

Obr. 15 Originální signál 50 EG silně ischemických cyklů a jejich kumulanty (Ischemické 2) 35

Obr. 16 Krabicové grafy koeficientů symetrie pro kumulant 3. řádu 39

Obr. 17 Krabicové grafy normalizované sumace pro kumulant 3. řádu 39

Obr. 18 Krabicové grafy variačních koeficientů pro kumulant 3. řádu 40

SEZNAM TABULEK

Tabulka 1 Korelační koeficienty a jejich směrodatné odchylky původních skupin cyklů a korespondujících kumulantů 37

Tabulka 2 Výsledky klasifikace na základě celých průběhů 43

Tabulka 3 Výsledky klasifikace na základě odvozených parametrů 44

ÚVOD

V dnešní době je nejčastější příčinou úmrtí a nemocnosti populace kardiovaskulární onemocnění. Nejčastěji používaným a také nejdostupnějším vyšetřením kardiovaskulárního systému je elektrokardiografie, díky které získáme záznam elektrické aktivity srdce. Vzniká tak křivka obsahující charakteristické vlny a kmity, která je pro stanovení diagnózy potřeba nějakým způsobem zhodnotit.

Samotnému hodnocení předchází předzpracování sejmutého elektrokardiogramu (EKG) a odstranění nežádoucích artefaktů a rušení vzniklých při snímání. K tomuto předzpracování je v dnešní době používáno mnoho různých algoritmů a další jsou stále vyvíjeny. Jedním z principů, na kterých může být předzpracování EKG signálu založeno, jsou i kumulanty vyšších řádů. Z předcházejících studií (např. z [9] a [17]) vyplývá, že kumulanty vyšších řádů jsou schopny díky svým vlastnostem potlačovat Gaussovský šum, snižovat rozdíly amplitud kmitů a vln a stejně tak i snižovat rozdíly způsobené časovým posunem cyklů. Kvůli těmto vlastnostem je velice užitečné kumulanty využít pro klasifikaci EKG cyklů do konkrétních skupin, která je prováděna i na základě morfologie EKG cyklů. Pro vysokou úspěšnost klasifikace je potřeba, aby variabilita cyklů v rámci jednotlivých skupin byla co nejmenší. A právě snížení variability EKG cyklů v rámci určité klasifikační skupiny je hlavním důvodem proč jsou kumulanty pro zpracování EKG záznamu využívané.

Cílem této bakalářské práce je ověření využitelnosti kumulantů vyšších řádů pro analýzu reálných signálů EKG. Signály, na kterých bude výzkum prováděn, byly získány z izolovaných králičích srdcí a jsou dostupné na Ústavu biomedicínského inženýrství VUT v Brně. Pomocí programového prostředí Matlab bude realizován výpočet kumulantů vyšších řádů a na nasimulovaných situacích budou ukázány a potvrzeny předpokládané vlastnosti kumulantů. V rámci práce bude také vytvořena databáze fyziologických a patologických EKG cyklů.

Nakonec budou cykly klasifikovány do odpovídajících skupin s využitím umělé neuronové sítě dvěma způsoby – klasifikace na základě původních cyklů a klasifikace pomocí statistických parametrů (příznaků) odvozených z kumulantů. Na základě úspěšnosti klasifikace bude diskutována využitelnost kumulantů vyšších řádů pro analýzu reálných dat a lékařskou praxi.

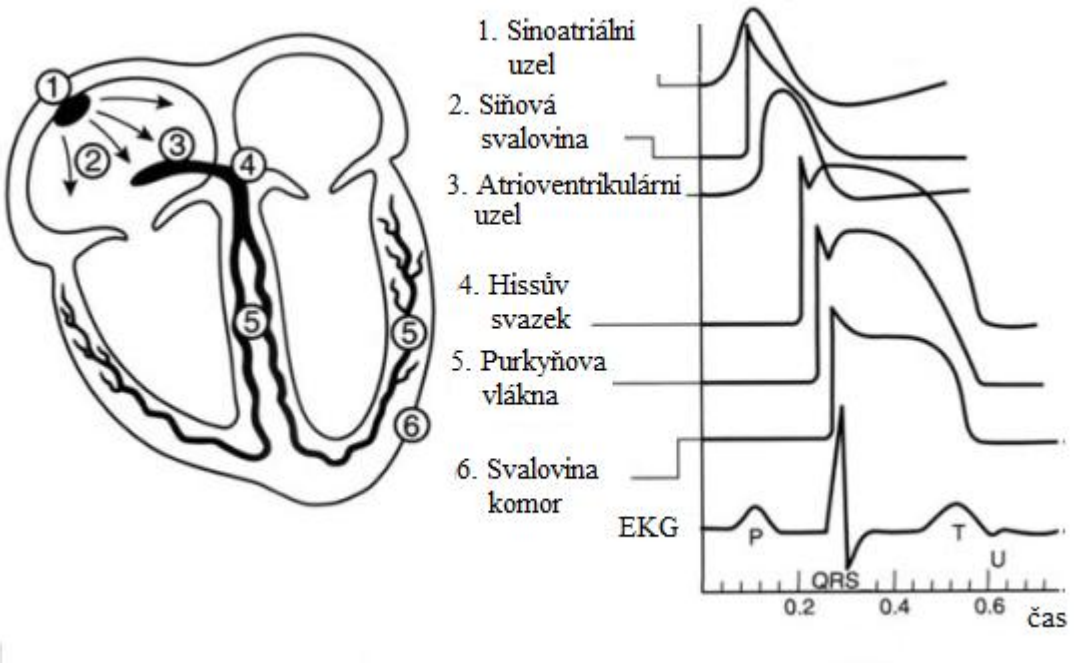
1 ELEKTROKARDIOGRAFIE

Elektrokardiografie je metoda, pomocí které zaznamenáváme elektrickou aktivitu srdce. Je to velmi často prováděné a cenné vyšetření, jelikož dokáže odhalit spoustu závažných změn v našem těle. Elektrická aktivita srdce se projevuje změnami napětí na povrchu těla. Tyto změny vznikají sumací všech potenciálů srdečních buněk (viz kapitola 1.1.). Tkáně lidského těla jsou díky velkému obsahu elektricky nabitých částic vodivé, a tak vznikají mezi různými částmi povrchu těla rozdílná napětí, která jsme schopni zaznamenat a následně zpracovat. Vznikne tak křivka - tzv. elektrokardiogram (EKG) - s charakteristickými vlnami a kmity jejichž morfologie, amplituda a načasování vypovídají o elektrickém vedení vzruchu v srdci, z čehož můžeme posoudit, zda srdce pracuje fyziologicky, nebo se jedná o určité srdeční onemocnění.

1.1 Šíření elektrického vzruchu srdcem

Šíření vzruchu je zajištěno *převodním systémem srdečním* (viz Obr. 1), což je soustava specializovaných svalových buněk schopná tvorby elektrických vzruchů a jejich vedení po srdečních oddílech. Excitace začíná zpravidla v *sinoatriálním uzlu*, což je shluk buněk se schopností spontánní depolarizace, kdy se v důsledku toku iontů přes plazmatickou membránu mění napětí na membráně a při překročení prahové hodnoty dojde ke vzniku *akčního potenciálu*. Ten se díky vodivým spojům (gap junctions) šíří jako vlna po srdeční svalovině. Ze svaloviny síní nemůže na svalovinu komor přímo přestoupit, jelikož jsou síně od komor elektricky izolovány. Jediné fyziologické spojení mezi nimi je přes *atrioventrikulární uzel*, kde dochází ke zpoždění šíření impulzu (atrioventrikulární uzel také rozhoduje o tom, zda vstupující impulz pošle dál, či nikoliv). Dále se vzruch šíří velkou rychlostí specializovanou vodivou tkání, která začíná úsekem zvaným *Hissův svazek*. Ten se v mezikomorovém septu dělí na *pravé* a *levé raménko Tawarovo*. Ve svalovině komor se tyto vlákna dále větví v menší vlákénka nazývané *Purkyňova vlákna*, která slouží k vlastnímu převodu vzruchu k buňkám *pracovního myokardu*. Takto se šířící impluz je fyziologický a rytmus jenž, díky němu vzniká, se nazývá *sinusový*. [4][6]

Kontrakce každého svalu jsou tedy spojené s elektrickými změnami zvanými *depolarizace*, což můžeme využít k detekci pomocí elektrod přiložených na povrch těla, kdy elektrickou aktivitu srdce měříme v podobě EKG (viz Obr. 1). Ačkoli má srdce jak známo čtyři oddíly, z hlediska měření elektrických potenciálů můžeme uvažovat pouze o dvou, protože pravá i levá síň se kontrahuje zároveň a také stah komor je společný.[7]



Obr. 1 Šíření elektrického vzruchu v srdci a vznik křivky EKG (převzato a upraveno z [14])

1.2 Způsoby záznamu EKG

Způsob, kterým se měří elektrická aktivita srdce je založen na připojení sady elektrod na lidské tělo (kůži). Rozdíl napěťových potenciálů mezi dvojicí elektrod se označuje jako svod. EKG se zaznamenává pomocí více svodů, které mohou být unipolární, či bipolární. Unipolární svody měří napětí mezi měřící a referenční elektrodou, která má nulový potenciál, svody bipolární zaznamenávají napětí mezi dvěma měřícími elektrodami umístěnými na různých místech. Mezi nejpoužívanější systémy záznamu patří *12-ti svodové EKG* a *ortogonální svodový systém*.

1.2.1 Standardní 12svodové EKG

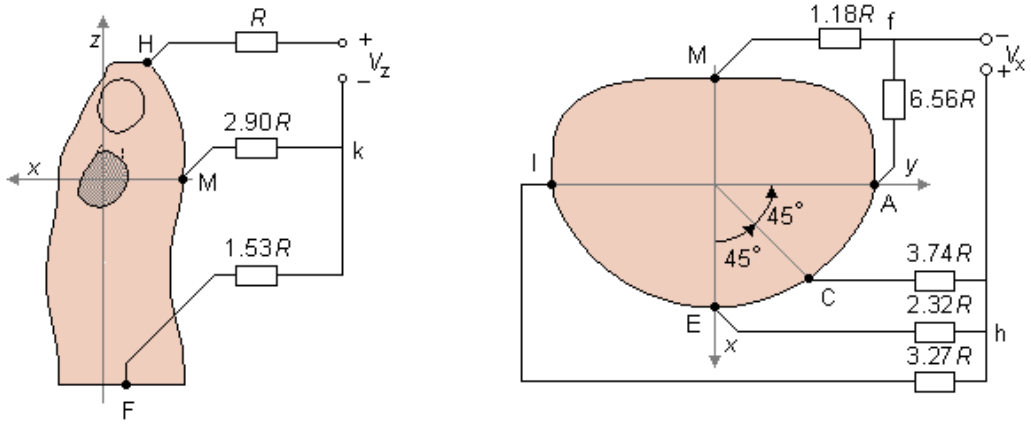
Dvanácti svodové EKG se skládá z kombinace tří různých typů svodů: bipolární končetinové svody, unipolární zesílené končetinové svody a unipolární hrudní svody. Všechny svody jsou zaznamenány pomocí 10ti elektrod umístěných na končetinách a hrudníku. *Bipolární končetinové* svody značené I, II, III se získají měřením napětí mezi pravou a levou rukou a levou nohou. *Unipolární zesílené končetinové* svody (aVR, aVF a aVL) využívají stejných elektrod jako svody bipolární ale porovnává se při tomto měření napětí mezi jedním ze svodů a průměrem ze zbývajících dvou. Další svody jsou *hrudní* (V1-V6), které jsou měřeny vzhledem k centrální (Wilsonově) svorce, která vznikne zprůměrování napětí ze všech končetinových svodů

Vlny EKG jsou v končetinových svodech menší a obsahují více rušení než v hrudních svodech, které mají amplitudu větší. To je dáno vzdáleností elektrod od srdce.[6]

1.2.2 Ortogonální svodový systém

Ortogonální svodové systémy odráží elektrickou aktivitu ve třech navzájem kolmých směrech X, Y a Z. V tomto svodovém systému tedy není popis EKG omezen pouze na analýzu jednotlivých svodů, ale můžeme získat další informaci skrze trojrozměrné zobrazení (tzv. vektorkardiogram).

Frankovými svody jsou nazývány opravené ortogonální svody, které jsou získány lineární kombinací sedmi elektrod umístěných na hrudníku, krku, zádech a levé noze. Jejich rozmístění a způsob zapojení je na obrázku 2. Výsledkem této kombinace jsou svody X, Y, Z, které zobrazují srdce v jednotlivých tělesných rovinách - frontální, horizontální a sagitální.[23]



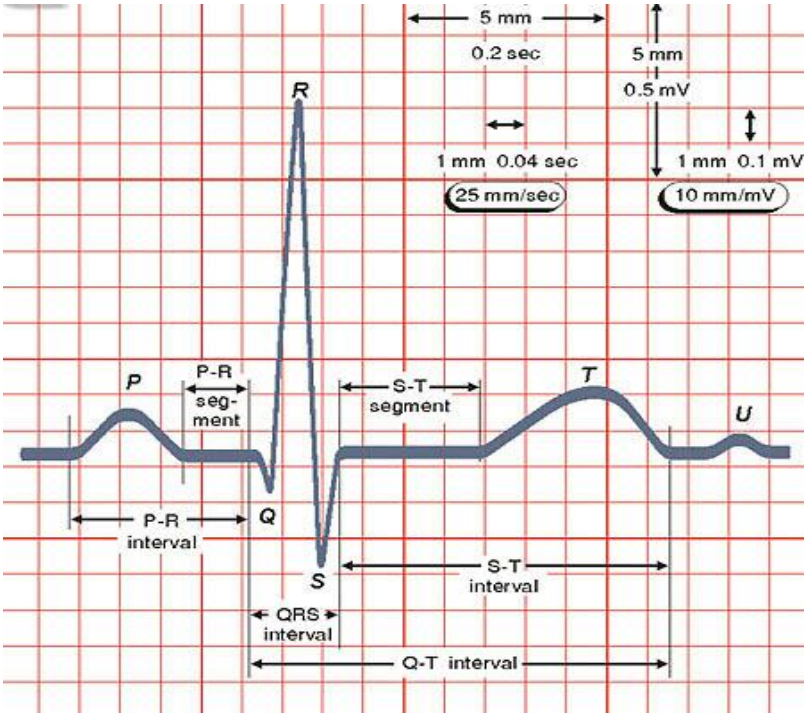
Obr. 2 Rozmístění a zapojení elektrod Frankova ortogonálního systému (převzato a upraveno z [23])

1.3 Popis EKG křivky

Popis jednotlivých vln a intervalů mezi nimi na EKG křivce je velmi důležitý pro následné zpracování signálu pomocí různých metod. EKG křivka se skládá z vlny P, komplexu QRS a vlny T. Fyziologický průběh EKG křivky s popisem vln je zobrazen na konci kapitoly na obrázku 3. Výchylka od *izoelektrické linie*, která se stanovuje z intervalu PQ, do kladných, či záporných hodnot vzniká dle toho, zda se elektrická vlna šíří směrem k elektrodě (výchylka nahoru), či směrem od ní (výchylka dolů). Důležitá je velikost a doba trvání jednotlivých vln:[6], [7], [22]

– První vlna na EKG křivce je *vlna P*, která odráží depolarizaci síní. Velikost P vlny je do 300 μV a její doba trvání kolem 100 ms, ve spektru EKG je zastoupena spektrálními složkami do 10-15 Hz. Může se stát, že v některých případech vlna P chybí (například při flutteru či fibrilaci síní, komorové nebo supraventrikulární tachykardii[5]). *Interval od začátku vlny P po začátek komplexu QRS* vypovídá o trvání šíření elektrického impulzu ze sinusového uzlu do srdečních komor.

- *Komplex QRS* odráží depolarizaci komor. První negativní kmit komplexu QRS komplexu se označuje jako *vlna Q*, první pozitivní jako *vlna R* a další negativní kmit následující po vlně R, jako *vlna S*. Fyziologicky QRS komplex trvá 80-100ms a maximální výchylka dosahuje velikosti 2-3 mV a z tohoto důvodu bývá ve většině programů pro analýzu EKG jako první detekován. Vyskytují se zde spektrální složky o vyšších frekvencích tj. 10-50 Hz. *Interval dvou po sobě jdoucích kmitů RR* vyjadřuje dobu trvání srdečního cyklu a je základním parametrem pro analýzu EKG (např. stanovení tepové frekvence a diagnostika arytmií).
- *Vlna T* je vyjádřením repolarizace myokardu komor, trvá kolem 200ms a vyskytuje se zhruba 300 ms za komplexem QRS, ovšem pozice vlny T je značně ovlivňována tepovou frekvencí a proto je její například její detekce hodně problematická.



Obr. 3 Fyziologická křivka EKG (převzato a upraveno z [22])

2 ÚVOD DO TEORIE KUMULANTŮ VYŠŠÍCH ŘÁDŮ

EKG může být popisováno mnoha způsoby, z nichž jedním jsou i kumulanty vyšších řádů, zejména řádu druhého, třetího a čtvrtého. Díky svým vlastnostem umožňují snížit variabilitu jednotlivých cyklů v rámci jedné konkrétní skupiny (např. normální cykly, ventrikulární či supraventrikulární extrasystoly, různé typy blokád aj.) v porovnání s původními cykly, nebo také potlačit Gaussovské rušení, jelikož vykazují vysoký poměr signál šum. V této kapitole bude popsáno matematické odvození kumulantů a jejich vlastnosti. Informace pro tuto část práce byly čerpány z [1], [2], [8], [17], [21], [26], [27].

2.1 Odvození kumulantů

Pro pochopení výpočtu kumulantů ze signálu, uvedeného v následující kapitole, je důležité nejprve objasnit pojem obecné momenty, jejichž pomocí se obvykle kumulanty definují. Momenty jsou pojmem z matematické statistiky a charakterizují pravděpodobností rozdělení náhodné veličiny. Pro zjednodušení bude níže uvedené odvození obecných momentů a pomocí nich i samotných kumulantů vysvětlováno pouze pro jednorozměrný proces.

Nejprve definujeme hustotu rozložení pravděpodobnosti náhodné proměnné X jako

$$\int_{-\infty}^{\infty} f(x)\,dx = 1, \tag{1}$$

kde f(x) je hustota pravděpodobnosti veličiny X. Pak pravděpodobnostní funkce spojité veličiny x se definuje následovně:

$$E(x) = \int_{-\infty}^{\infty} g^x f(x)\,dx. \tag{2}$$

Pokud nyní položíme v rovnici (2) $g^x=e^t$, získáme momentovou vytvořující funkci $M_x(t)$ náhodné veličiny X :

$$M_X(t) = E(e^x) = \int_{-\infty}^{\infty} e^{tx} f(x)\,dx. \tag{3}$$

Pro diskrétní náhodnou veličinu X je momentová vytvořující funkce definována takto:

$$M_X(t) = E(e^{tx}) \sum_x e^{tx} P(x), \tag{4}$$

kde P(x) = P(X=x) je pravděpodobnostní funkce veličiny X.

Vytvořující funkci $M_X(t)$ lze vyjádřit pomocí Taylorovy řady, jejíž definice je:

$$s(x) = \sum_{n=0}^{\infty} \frac{f^n(a)}{n!}(x-a)^n = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots. \quad (5)$$

Z výše uvedeného je nám známo, že $M_X(t) = E(e^{tX})$. Pokud položíme v Taylorově polynomu a=0 a zároveň využijme znalosti $e^0=1$ a získáme rozvoj momentové vytvořující funkce v Taylorův polynom

$$\begin{aligned} M_X(t) &= E\left[e^0 + \frac{X}{1!}(t-0) + \frac{X''}{2!}(t-0)^2 + \frac{X'''}{3!}(t-0)^3 + \cdots..\right] \\ &= E\left[e^0 + \frac{X\cdot t}{1!} + \frac{X''\cdot t^2}{2!} + \frac{X'''\cdot t^3}{3!} + \cdots..\right] \\ &= E(1) + \frac{E(X)\cdot t}{1!} + \frac{E(X'')\cdot t^2}{2!} + \frac{E(X''')\cdot t^3}{3!} + \cdots. \end{aligned} \quad (6)$$

Vztahu (5) využijeme nyní při důkazu toho, že momenty dostaneme jako derivaci charakteristické vytvořující funkce $M_x(t)$ v nule. Zderivujeme ji a položíme t=0

$$\begin{aligned} &\frac{dt}{t}\left\{E(1) + \frac{E(X)\cdot t}{1!} + \frac{E(X'')\cdot t^2}{2!} + \frac{E(X''')\cdot t^3}{3!} + \cdots..\right\}\Big|_{t='0} \\ &= 0 + \frac{E(X)}{1!} + \frac{2\cdot E(X'')\cdot t}{2!} + \frac{3\cdot E(X''')\cdot t^2}{3!} + \cdots\Big|_{t=0} \\ &= \frac{E(X)}{1!} + \frac{2\cdot E(X'')\cdot 0}{2!} + \frac{3\cdot E(X''')\cdot 0^2}{3!} = E(X) = m_1 \end{aligned} \quad (7)$$

Dostali jsme první obecný moment m_1. Pokud bychom nyní provedli druhou derivaci této funkce získáme druhý obecný moment. Z toho již plyne základní význam momentové vytvořující funkce a to že její n-tá derivace v bodě 0 dává n-tý obecný moment m_n.

Pokud nyní momentovou generující funkci zlogaritmujeme, získáme kumulanty generující funkci $g_x(t)$ [17]

$$\begin{aligned} g(t) = \log\big(E(e^{tX})\big) &= -\sum_{n=1}^{\infty}\frac{1}{n}(1-E(e^{tX}))^n = -\sum_{n=1}^{\infty}\frac{1}{n} - \left(\sum_{m=1}^{\infty} m_n \frac{t^m}{m!}\right)^n \\ &= m_1 + (m_1 - m_1^2)\frac{t^2}{2!} + (m_3 - 3m_2m_1 - 3m_2^3)\frac{t^3}{3!} + \cdots. \end{aligned} \quad (8)$$

A nyní stejným způsobem jako jsme výše dostali momenty, dostaneme konečně kumulanty, čili n-tá derivace kumulanty generující funkce g(t) v bodě nula nám dá n-tý kumulant. V této práci budou využity kumulanty 1-4. řádu, proto jsou níže uvedeny jejich výsledné definice, které jsou získány z rovnice (8) provedením výše popsaných operací:

$$c_1 = m_1, \quad (9)$$

$$c_2 = m_2 - m_1^2, \tag{10}$$

$$c_3 = m_3 - 3m_2m_1 - 3m_1^3, \tag{11}$$

$$c_4 = m_4 - 4m_3m_1 - 3m_2^2 + 12m_2{m_1}^2 - 6{m_1}^4. \tag{12}$$

2.2 Tvorba kumulantů ze signálu v praxi

Kumulanty (tzn. „autokumulanty“) ze signálu x(n) pro n = $\langle 0; N-1 \rangle$, který je reálný, diskrétní a náhodný a má nulovou střední hodnotu, vypočítáme tak, že budeme mezi sebou násobit vzorky tohoto signálu, posunuté vůči sobě o určitý časový okamžik (tj. určitý počet vzorků):

$$c_{2\mathrm{x}}(k) = \mathrm{m}_{2x} = E[x(n)x(n+k)] = \frac{1}{N3}\sum_{n=N1}^{N2} x^*(n)x(n+k), \tag{13}$$

$$c_{3\mathrm{x}}(k,l) = m_{3x} = E[x(n)x(n+k)x(n+l)] = \frac{1}{N3}\sum_{n=N1}^{N2} x^*(n)x(n+k)\,x(n+l) \tag{14}$$

$$c_{4\mathrm{x}}(k,l,m) = E[x(n)x(n+k)x(n+l)x(n+m)] - c_{2\mathrm{x}}(k)c_{2\mathrm{x}}(m-l) -$$
$$c_{2\mathrm{x}}(l)c_{2\mathrm{x}}(m-k) - c_{2\mathrm{x}}(m)c_{2\mathrm{x}}(k-l) = \frac{1}{N3}\sum_{n=N1}^{N2} x^*(n)x(n+k)\,x(n+l)x^*(n+m) -$$
$$c_{2x}(k)c_{2x}(l-m) - c_{2x}(l)c_{2x}(k-m) - m_{2x}(m)m_{2x}(l-k), \tag{15}$$

kde k, l, m jsou časové okamžiky (posunutí signálu), x*(n) je komplexně sdružená posloupnost vstupního signálu, N1 a N2 nabývají hodnot rozmezí 0 až N-1 [19]. Pro výpočet „neposunutého odhadu“ (angl. unbiased) kumulantů se N3 nastavuje na hodnotu, která odpovídá délce (počtu vzorků) posloupnosti aktuálně využívané pro výpočet kumulantu, při výpočtu posunutého (angl. biased) odhadu se zadává hodnota menší, než je délka signálu.

Pro nulové hodnoty k, l a m reprezentují normalizované hodnoty kumulantů známé charakteristiky rozdělení náhodných veličin a to: kumulant 2. řádu reprezentuje rozptyl, kumulant 3. řádu – koeficient šikmosti a kumulant 4. řádu - koeficient špičatosti. [9]

2.3 Základní vlastnosti kumulantů

Nejdůležitější vlastnosti kumulantů jsou následující [8], [22]:

1. Jestliže x(n) je Gaussovský proces, potom k-tý kumulant veličiny x(n) pro k > 2 bude nulový. A tedy veškeré informace budou obsaženy v prvním a druhém kumulantu. Této vlastnosti můžeme využít pro potlačení Gaussovského šumu neznámé střední hodnoty.
2. Jestliže x(n) má symetrickou distribuci kolem nuly, potom každý k-tý kumulant x(n) pro k= 0,3,5,7 bude nulový. Toho se využívá například pro zjištění tvaru signálu.
3. Jestliže x(n) má nulovou střední hodnotu, potom k-tý kumulant bude roven k-tému obecnému momentu pro k <= 3. Díky tomu lze velice zjednodušit výpočet.

4. Jestliže z(n) = x(n) + y(n), kde x(n) a y(n) je stacionární a statisticky nezávislý proces, potom se kumulant veličiny z(n) bude rovnat součtu kumulatnů veličin x(n) a y(n). Toho by se dalo využít při detekci aditivních směsí v signálu. Tato vlastnost neplatí pro momenty.

Studie, o kterých mluví kapitola č. 3, ukazují, že tyto vlastností je možno využít pro zpracování EKG signálu. V kapitole č. 5 jsou tyto vlastnosti otestovány a ověřeny na simulovaných a naměřených datech.

3 VYUŽITÍ KUMULANTŮ V SOUČASNÉ PRAXI

Počítačově asistovaná detekce srdeční arytmie hraje důležitou roli v odhalení arytmie vůbec a poté i ve stanovení konkrétního onemocnění a následné léčbě. V současné praxi je pro tuto problematiku využíváno mnoho různých způsobů a jejich kombinací, jak správně klasifikovat morfologicky různé EKG cykly a tím zjistit konkrétní typ arytmie. Stále ovšem klasifikace není stoprocentně úspěšná. Právě pro zlepšení této úspěšnosti se využívá také kumulantů vyšších řádů. Díky jejich vlastnostem potlačují variabilitu EKG cyklů v rámci jedné skupiny v porovnání s původními cykly. Po vypočítání kumulantů ze signálu získáme cykly, které jsou pro konkrétní skupinu charakteristické, což umožňuje jejich následnou snadnější klasifikaci.

V současné literatuře nacházíme, že kumulanty nebývají často při zpracování EKG používány samostatně, ale jsou kombinovány s jinými principy. Jedním z nich je například Hermitain bacis function. Kombinace těchto metod bylo využito například v [9]. Nejprve byly vypočítány kumulanty z EKG cyklů a následně namodelovány lineární kombinace Hermitain Basic Function. Z každého cyklu byly namodelovány 3 reprezentativní vektory a to na základě kumulantů druhého, třetího a čtvrtého řádu. Autoři třídili EKG cykly do pěti klasifikačních skupin (normální cykly, supraventrikulární extrasystoly, ventrikulární extrasystoly, blokáda pravého a levého Tawarova raménka) použitím klasifikační metody nejbližšího souseda (angl. Nearest neighbour, k-NN). Po definování reprezentativních cyklů pro každou klasifikační skupinu byla metoda testována. Bylo použito 5621 cyklů pro trénování klasifikátoru a 3746 jich bylo využito pro testování úspěšnosti klasifikace. Pro každý cyklus byly tedy porovnávány 3 vektory různých řádů kumulantů a pro klasifikaci bylo třeba, aby nejméně dva byly zařazeny do stejné skupiny. Pokud byl každý zařazen do rozdílné skupiny, klasifikace byla provedena dle kumulantu 3. řádu, jelikož podle předchozího testování vykazoval největší spolehlivost. Hodnota specificity detekce arytmií byla 99.67% a sensitivity 98.66%, což ukazuje opravdu velkou sílu a spolehlivost algoritmu. Vysoké citlivosti bylo dosaženo díky tomu, že metoda velice dobře potlačovala Gaussovský šum a také časový posun a rozdíl amplitudy cyklů v rámci jedné klasifikační skupiny.

Další studií, kde bylo použito podobného postupu jako výše, je [19]. Hodnoty senzitivity a specificity zde byly 75 % respektive 76 %. Nižší citlivost než v předchozím případě byla způsobena zejména tím, že klasifikace byla prováděna pouze na základě jednoho reprezentativního vektoru kumulantu každého cyklu a ne pomocí kombinace 3 vektorů. Ve studii byla ovšem metoda porovnávána s jinou identickou, která posuzovala tytéž EKG cykly a vykazovala hodnoty senzitivity a specificity pouze 46 % a 45%, což opět vypovídá o velké spolehlivosti. Stejných metod bylo využito i v [18].

Dále jsou často kumulanty pro klasifikaci EKG využívány v kombinaci s předzpracováním signálu pomocí vlnkové transformace (angl. wavelet transform), což je metoda, která je vhodná pro v čase proměnné signály, kterým signál EKG je. Umožňuje pomocí předem určené "vlnky" rozložit signál a zjistit tak zastoupení frekvenčních složek. Kombinace vlnkové transformace a kumulantů vyšších řádů autoři využili např. v [24]. Nejprve byla provedena 5ti násobná vlnková transformace, následně vypočítány kumulanty a pro klasifikaci byla využita dopředná vícevrstvá neuronová síť. Metoda byla testována na základě 4 různých umělých způsobů zarušení a 2 různých způsobů výběru vzorků. Úspěšnost klasifikace (tedy procento cyklů správně zařazených do odpovídajících skupin) byla v průměru 97,5% a bylo dokázáno, že metoda zlepšuje citlivost klasifikace. Podobných postupů bylo využíváno také v [15], [16].

Ve studii [27] bylo kumulantů využito ve spojení se zajímavým způsobem klasifikace cyklů pomocí neuronové sítě, kdy nebyli vstupem sítě samotné průběhy cyklů, ale pouze 30 statistických parametrů z nich odvozených. Úspěšnost klasifikace se pohybovala kolem 98 %. Tento jiný způsob klasifikace je výhodný pro menší technické nároky z důvodu menšího počtu vstupních vzorků signálu.

Kumulantů EKG cyklů bylo využito i v jinak než pro klasifikaci arytmií. Jednalo se o identifikaci konkrétních osob (např. pro vstup do objektu) pomocí EKG záznamu [1]. Nejprve bylo nahráno několik záznamů od každé osoby a pomocí kumulantů druhých řádu (tzn. autokorelace) byli vytvořeny vzorky pro následné porovnávání pomocí korelačního koeficientu se vzorky sejmutými od osoby pokoušející se o identifikaci.

Jak známo, analýzu EKG nemusíme provádět pouze v oblasti časové, ale i v oblasti spektrální. I v této oblasti je možno uplatit kumulanty – tvorbou bispekter. Tohoto způsobu bylo využito v [10].

Z předchozích odstavců vyplývá, že kumulanty v současné době nacházejí uplatnění zejména v oblasti klasifikace srdečních arytmií. Velice často nejsou kumulanty využívány samostatně, ale jsou kombinovány s jinými metodami zpracování signálu a to např. s vlnkovou transformací nebo s Hermitain basic function. Ve všech zmiňovaných výzkumech využití kumulantů zpřesnilo klasifikaci srdečních cyklů do konkrétních skupin, zejména díky tomu, že dokázali potlačovat Gaussovský šum a snižovat variabilitu srdečních cyklů v rámci jedné skupiny.

4 REALIZACE VÝPOČTU KUMULANTŮ VYŠŠÍCH ŘÁDŮ V PROSTŘEDÍ MATLAB

Matlab je programové prostředí určené pro vědeckotechnické účely, tvorbu algoritmů, analýzu dat, tvorbu aplikací včetně grafického rozhraní atd. Pro realizaci výpočtů v této práci bylo využito tohoto programovacího prostředí. Pro výpočty byl použit osobní počítač s procesorem Celeron(R) Dual-Core CPU T3300 @ 2.00 GHz a operačním systémem Microsoft Windows 7 Home Premium.

4.1 Předzpracování elektrokardiogramu

Signály, které jsou zpracovávány v této práci, jsou dostupné na ústavu Biomedicínského inženýrství VUT v Brně. Tyto elektrogramy (EG) byly získány z izolovaných králičích srdcí v průběhu experimentu s opakovanou fází ischemie a reperfuze. Podrobnější popis metody lze najit v [11]. Vzhledem k tomu, že signály byly získány z izolovaných srdcí, mluvíme o těchto záznamech jako o elektrogramech a ne jako je tomu obvyklé při standartním snímání rozdílu elektrických potenciálu o elektrokardiogramech.

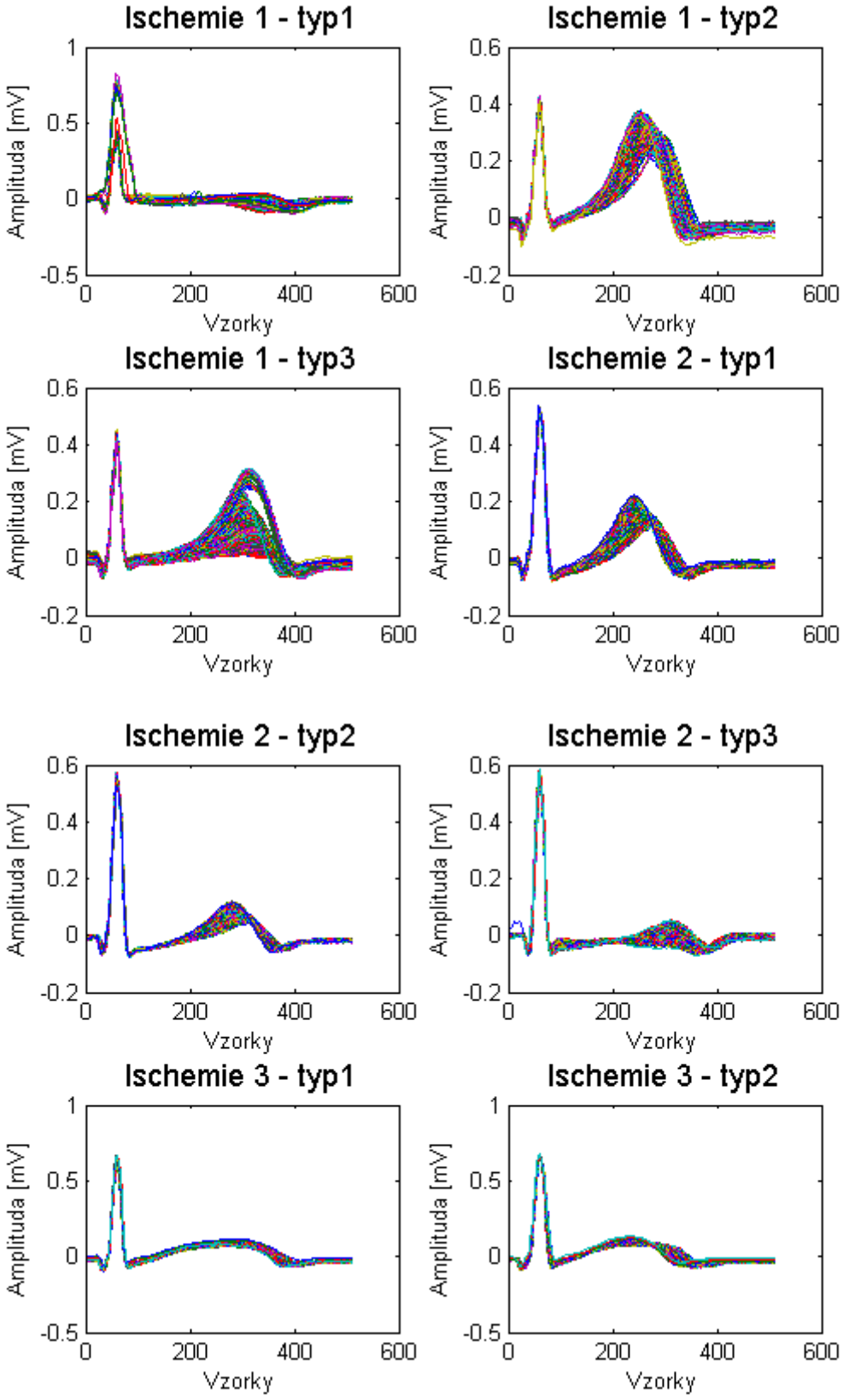
Z naměřených elektrogramů byly vybrány reprezentativní úseky, tedy bez artefaktů a ty byly rozděleny dle fází experimentu. Úseky signálu byly filtrovány pomocí Lynnova filtru typu horní propust s mezní frekvenci 0,5 Hz pro odstranění kolísání nulové izolinie, ke kterému docházelo především u svodu III. Software, který byl pro tuto filtraci využit je [12]. Poté byly detekovány maxima R vln pomocí detektoru [25] pracujícího na základně vlnkové transformace. Nakonec byly vybrány pouze úseky QRST (tedy bez vlny P) jako úseky signálu 59 vzorků před detekovanou R vlnou až 450 vzorků po R vlně. Výsledné cykly EKG měly vždy délku 510 vzorků. Takto získané cykly byly uloženy do cell buňky o rozměrech 7 řádků zastupujících fáze experimentu a 4 sloupců, kdy v prvním z nich je vždy uveden název fáze a v dalších třech jsou uloženy svody X, Y, Z.

Cykly, rozdělené dle fází experimentu, byly vždy v rámci této fáze ještě roztříděny na určitý počet typů cyklů podle morfologie (počet typů byl v jednotlivých fázích různý dle variability cyklů v rámci fáze). Byla tak vytvořena databáze cyklů (opět cell buňka) určitých morfologických typů EG cyklů. Roztřízené skupiny cyklů z databáze jsou zobrazeny na obrázcích 4, 5 a 6. Vždy je zobrazen svod X. To, že cykly pocházejí např. z ischemické fáze experimentu, zároveň neznamená, že jejich morfologické znaky odpovídají znakům charakteristickými pro cykly značící ischemii myokardu.

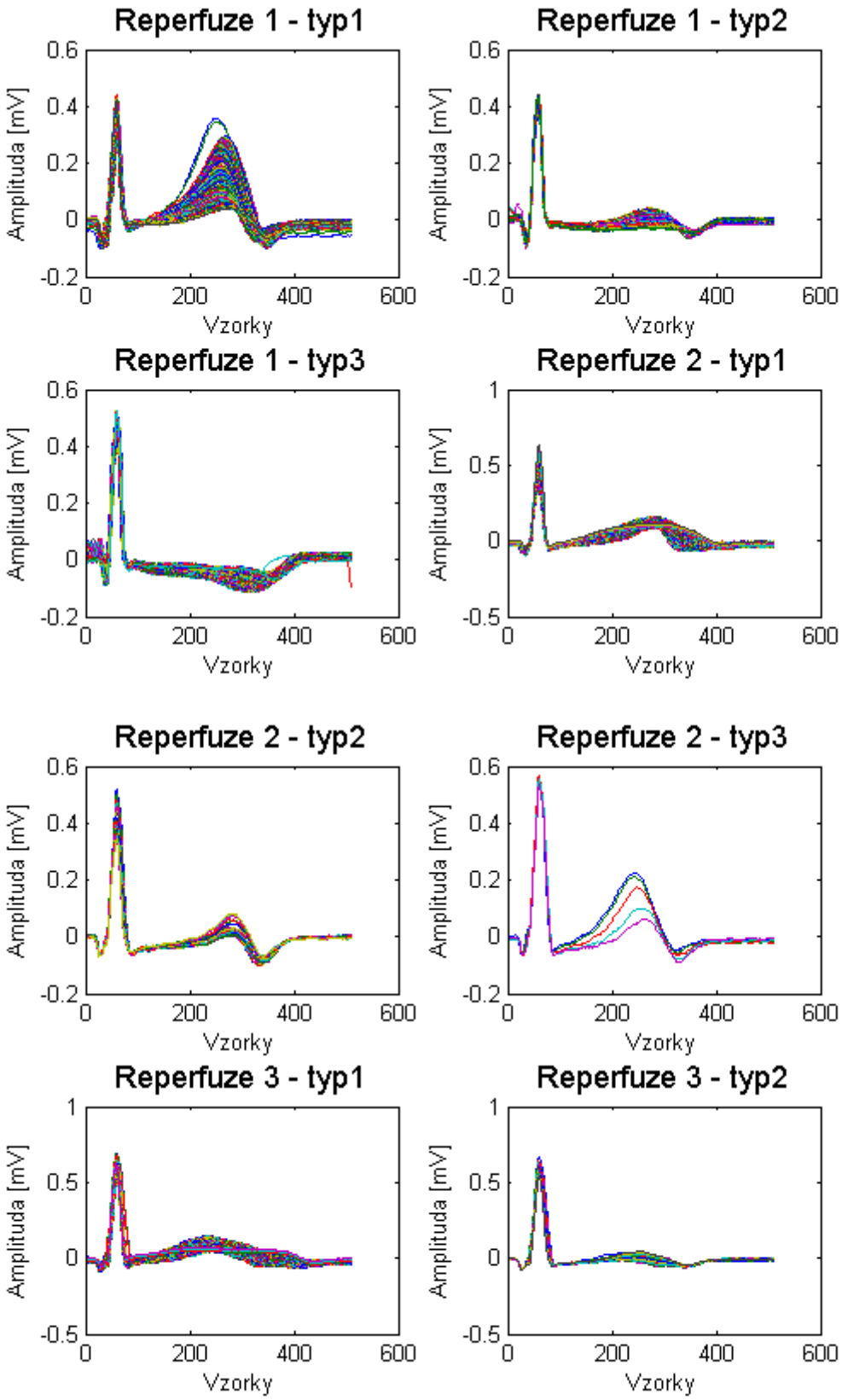
Cykly byly tříděny na základě morfologických znaků, kterými byli:

- ST úsek (s elevací a bez elevace),
- polarita vlny T (pozitivní či negativní),

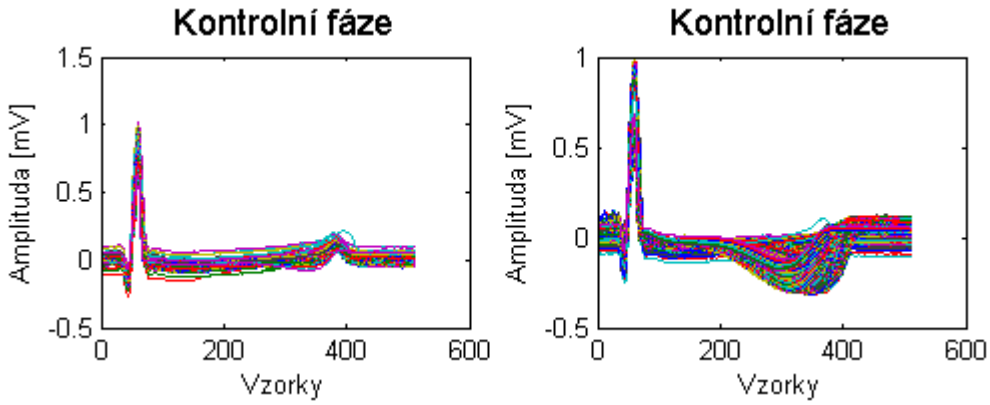
- tvar vlny T (výrazná a tedy s velkou amplitudou, či s amplitudou nízkou),
- přítomnosti či nepřítomnosti zákmitu pod izolinii těsně za vlnou T.



Obr. 4 Cykly z ischemických fázích roztříděné dle morfologie



Obr. 5 Cykly z fází reperfuze roztříděné dle morfologie

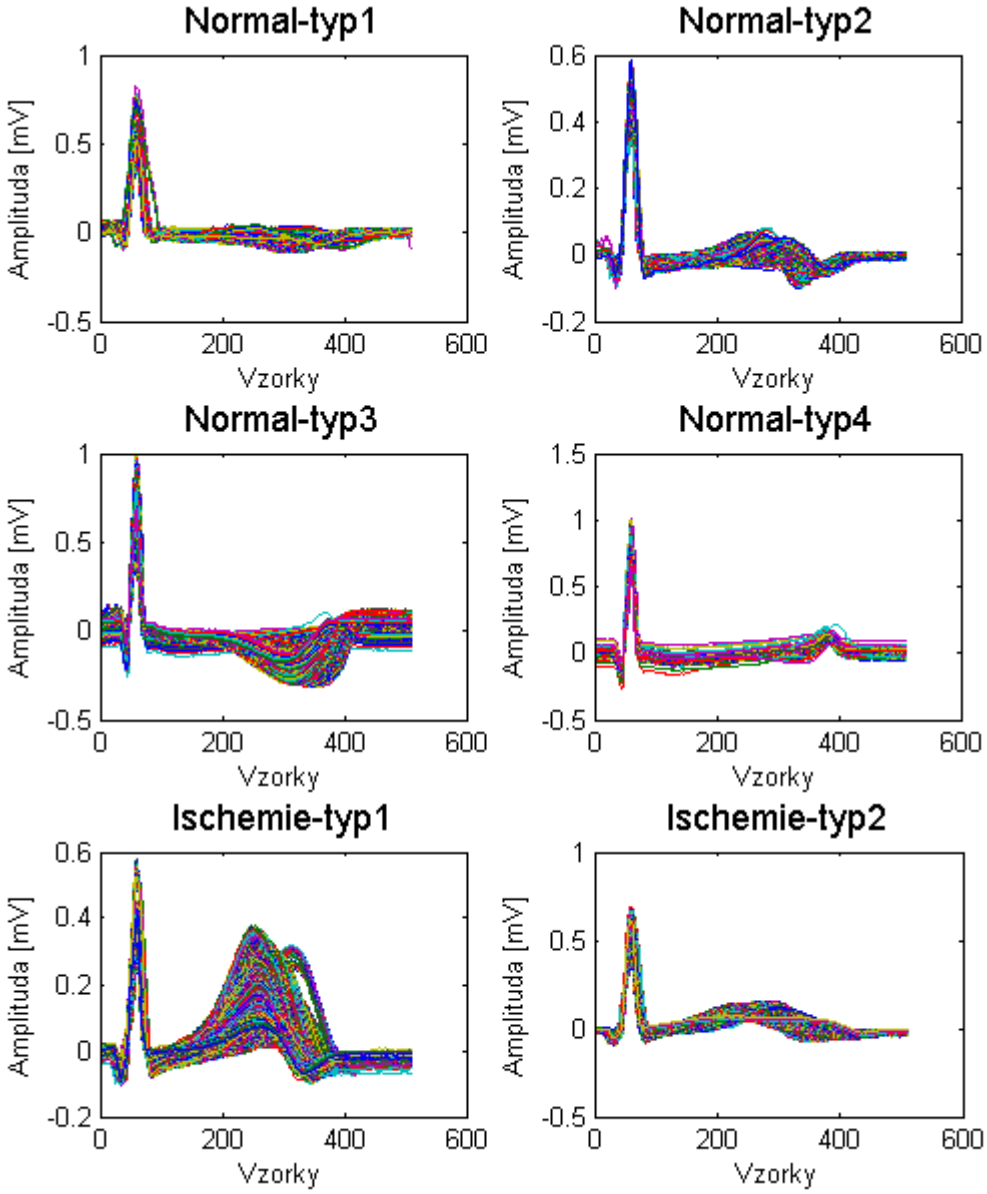


Obr. 6 Cykly z kontrolní fáze roztříděné dle morfologie

Jak bylo předpokládáno, v rámci jednotlivých fází se vyskytovali typy cyklů, které si byli velice podobné s cykly, vyskytujícími se v jiné fázi. Je to z toho důvodu, že například na začátku ischemické fáze se ischemie na srdečním svalu plně neprojevila a srdce ještě pracovalo jako při dostatečném okysličení a naopak na začátku fáze reperfuze srdce ještě nemělo dostatek kyslíku potřebného pro fyziologickou funkci a proto si jsou morfologicky podobné i cykly nepocházející ze stejných fází.

Po sloučení typů cyklů, které si byly navzájem podobné, vzniklo výsledných 6 typů zobrazených na obr. 7. Názvy těchto skupin typů již odpovídají tomu, do jaké klasifikační skupiny ve skutečnosti patří. Z toho důvodů je jich tedy možné využít jako směrodatná vstupní data pro případné zhodnocení úspěšnosti klasifikace. Zařazení typů cyklů do těchto klasifikačních skupin (normální, ischemické) bylo stanoveno na základě charakteristických morfologických znaků, které jsou daným typům vlastní. Znalosti využité k tomuto zařazení byly získány z [5] a [6] nebo z vlastních pracovních zkušeností nabytých při hodnocení EKG holterů.

V rámci práce byl použit i signál získaný stejným způsobem od jiného zvířete. Bylo to učiněno z důvodu výskytu více různých typů cyklů, než obsahoval výše zobrazený a popisovaný signál.



Obr. 7 Normální cykly ze všech fází roztříděné a sloučené na základě morfologie

4.2 Popis použitých funkcí

K výpočtu kumulantu z EKG cyklu bylo využito funkce s názvem *cumest*. Hlavička funkce je:

function y_cum = cumest(y,norder,maxlag,nsamp,overlap,flag,k1,k2).

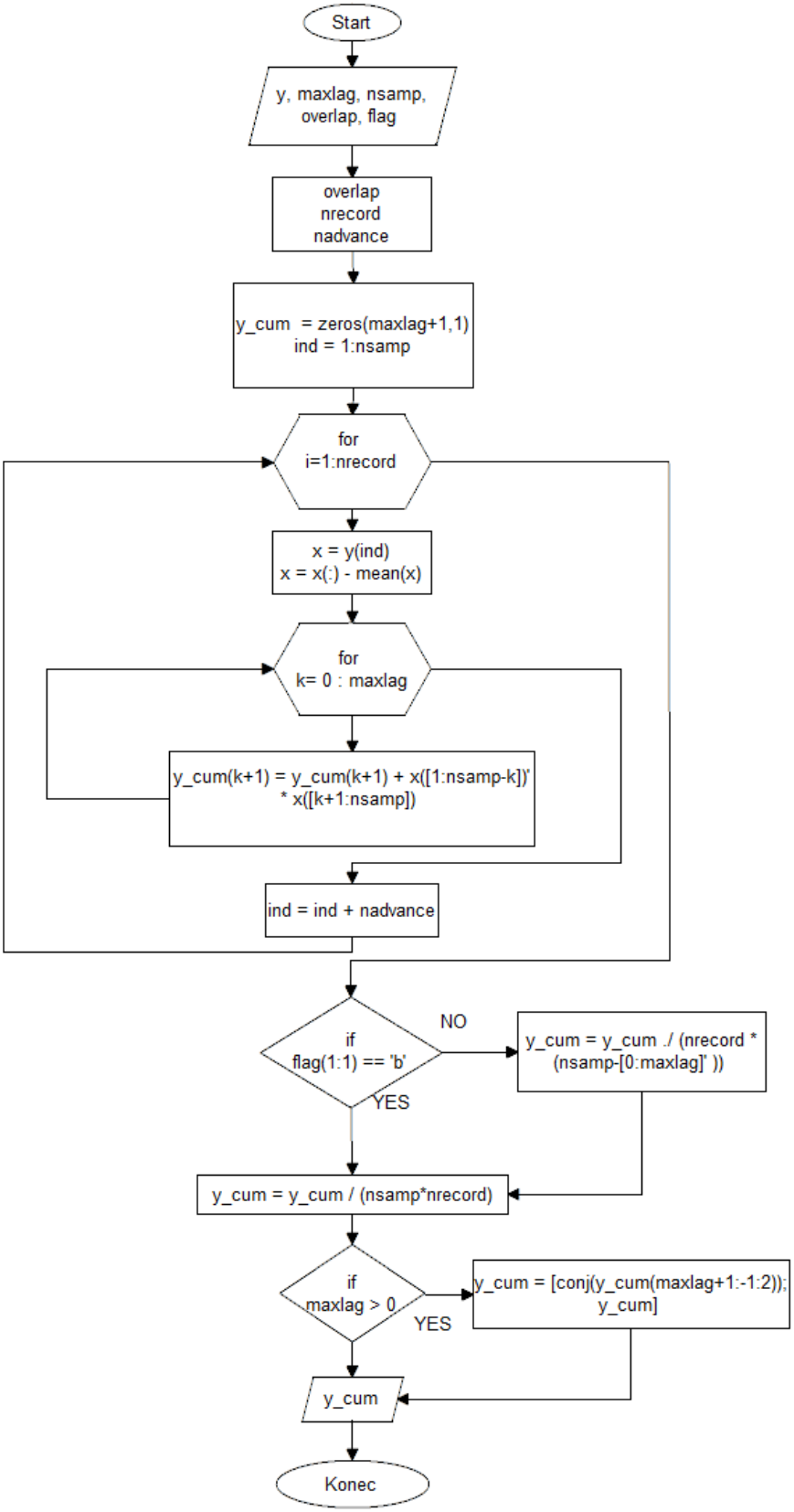
Vstupními proměnnými funkce cumest jsou:

– *y* – vstupní vektor hodnot pro výpočet kumulantu (zde tedy jeden EKG cyklus o délce 510 vzorků),

- *norder* – výběr řádu kumulantu, jenž bude počítán, může nabývat hodnot 2, 3 nebo 4, přednastavena je hodnota 2, čili druhý řád, dle zvoleného řádu bude pro výpočet použita příslušná funkce pro výpočet konkrétního řádu kumulantu tj. funkce *cum2est, cum3est, cum4est* (další vstupní proměnné budou tedy využity zvolenou funkcí);
- *maxlag* – určuje maximální časový posun pro výpočet kumulantu;
- *nsamp* – počet vzorků segmentu, tedy z kolika vzorků, pokud ne najednou z celého vektoru, budou počítány kumulanty, pokud chceme počítat kumulant z celého vektoru naráz, bude mít tato vstupní proměnná hodnotu délky vektoru, ze kterého bude kumulant počítán;
- *overlap* – udává procento překrytí segmentů (které jsou určeny pomocí parametru *nsamp*), přednastaveno je 0, možné hodnoty této proměnné leží v intervalu < 0, 99 >;
- *flag* – přípustné hodnoty této proměnné jsou: *'biased'* a *'unbiased'*, defaultně je nastaveno *'biased',* možnost *'biased'* značí, že se jedná o posunutý odhad, možnost *'unbiased'* znamená odhad neposunutý;
- *k1* – časový posun, nastavuje se pouze v případě hodnoty proměnné *norder* 3 nebo 4
- *k2* – časový posun, nastavuje se pouze v případě hodnoty proměnné *norder* 4.

Výstupní proměnnou je *y_cum* což je při výpočtu kumulantu druhého řádu vektor hodnot o délce 2*N-1 (N je délka původního signálu), v případě výpočtu kumulantu druhého řádu se jedná o 2D matici respektive o 3D matici v případě kumulantu řádu čtvrtého.

Samotná funkce *cumest* je schopna realizovat výpočet kumulantu z jednoho cyklu. V této práci bylo potřeba vypočítat kumulanty z větší skupiny cyklů. Proto byl výpočet realizován ve *for* cyklu. Při výpočtu kumulantů 3. a 4. řádu ze skupiny cyklů najednou bylo využito dvojitého *for* cyklu, kdy první slouží k postupnému vybrání jednotlivých cyklů EKG z matice hodnot konkrétního svodu (nejprve jsou vždy načteny jednotlivé svody a uloženy do odpovídajících proměnných). V druhém *for* cyklu jsou nastavovány mezní hodnoty *k1* a *k2*, podle kterých se následně realizuje vzájemný posun signálu při samotném výpočtu kumulantu, který je právě v tomto cyklu pomocí funkce *cumest* realizován. Při výpočtu kumulantu řádu druhého probíhá výpočet v prvním z for cyklů. Dle hodnoty vstupní proměnné norder jsou následně volány funkce s názvy *cum2est, cum3est* a *cum4est*, které realizují výpočet kumulantů příslušných řádů pomocí operací uvedených v kapitole 2.2. Funkce *cum2est* tedy realizuje výpočet kumulantu druhého řádu operací popsanou rovnicí (12), funkce *cum3est* realizuje rovnici (13) a konečně funkce *cum4est* rovnici (14). Blokové schéma samotné funkce *cum2est* je uvedeno a popsáno níže.



Obr. 8 Blokové schéma funkce *cum2est* pro výpočet kumulantu druhého řádu

Popis blokového schématu výpočtu kumulantu 3. řádu a ilustrativní příklad výpočtu z konkrétního signálu y = [1 7 4] :

1. Načítání vstupních proměnných.

```
cumest([1 7 4],2, nsamp-1, size(y,2),0, 'unbiased')
```

2. Na základě zadaných vstupních proměnných vytvořeny pomocné proměnné, které budou dále využívány a to: *overlap* - procento překrytí sousedních segmentů (při výpočtu kumulantu z úseku rozděleného na jednotlivé segmenty) je přepočítáno na počet vzorků, *nrecord* – udává počet segmentů, ze kterých se počítá hodnota kumulantu, tj. kolikrát proběhne první (vnější) cyklus for (na základě hodnoty *nsamp*, pokud je rovna délce vektoru y, proběhne cyklus jen jednou), *nadvance*- hodnota, která bude přičtena k indexu *ind*, tj. výsledek určuje vzorky, ze kterých bude kumulant počítán v příštím for cyklu.

```
overlap  = fix(overlap/100 * nsamp); => overlap = 0;
nrecord  = fix( (N - overlap)/(nsamp - overlap) ); => nrecord=1;
nadvance = nsamp - overlap; => nadvance=3;
```

3. Příprava nulového vektoru, do něhož budou ukládány hodnoty budoucího kumulantu a indexu *ind*, který určuje, z kterých vzorků signálu *y* bude část kumulantu nejprve počítána

```
y_cum = [0;0;0]
ind = [1,2,3]
```

4. For cyklus, ve kterém se z úseku signálu nejdříve odečítá střední hodnota (protože algoritmus je založen na zjednodušeném výpočtu kumulantu ze signalu s nulovou střední hodnotou, viz kap. 2.2) a z takto předzpracovaného signálu dojde v následujícím vnořeném for cyklu k vypočítání vzorků kumulantu.

```
y_cum = [-3;3;0]
```

5. V tomto vnořeném *for* cyklu je realizováno násobení jednotlivých hodnot signálu mezi sebou postupně posouvaných o jednotlivé vzorky. Násobí se převrácený vektor signálu s identickým, ale transponovaným vektorem.

```
y_cum = [18;-9;0]
```

6. Pokud je vstupní proměnná *flag* rovna *biased*, je realizován výpočet posunutého odhadu, v opačném případě je proveden výpočet odhadu neposunutého.

```
Neposunutý odhad: y_cum = [6;-4.5;0]
Posunutý odhad: y_cum = [6;-3;0]
```

7. Pokud je hodnota *maxlag* > 0, dojde k doplnění hodnot *y_cum(end:2)* před *y_cum*.

```
Neposunutý odhad: y_cum = [0;-4.5;6;-4.5;0]
Posunutý odhad: y_cum = [0;-3;6;-3;0]
```

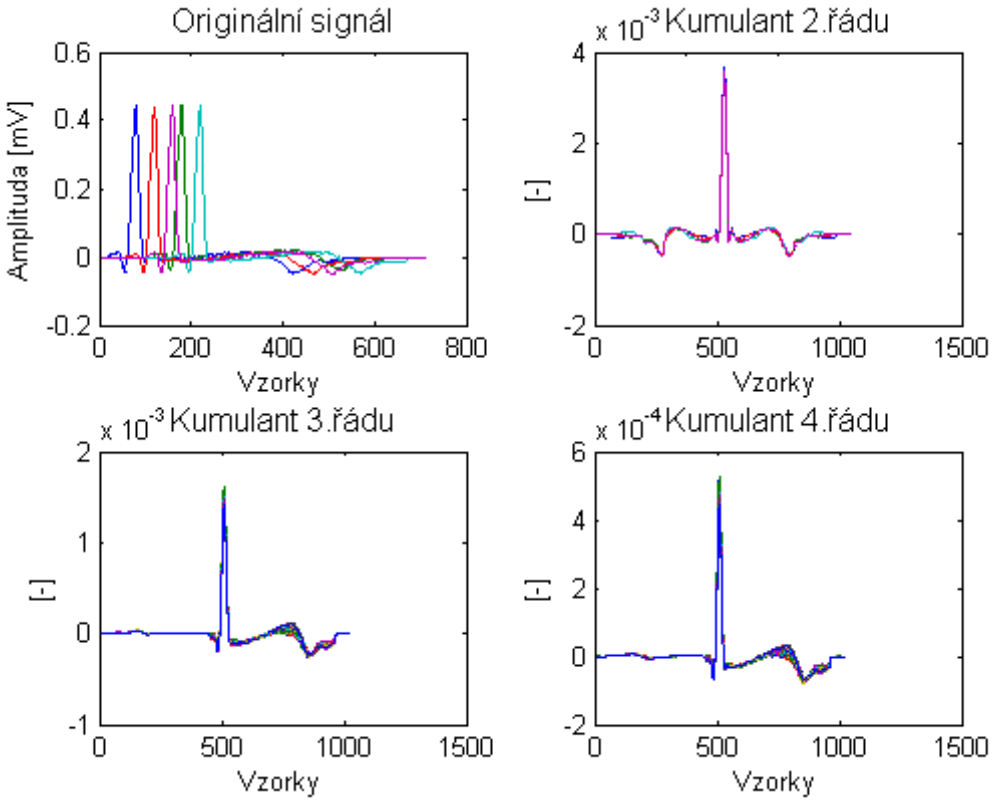
V případě výpočtu kumulantu 3. řádu probíhá výpočet obdobným způsobem, je ovšem přidána proměnná *lag*, jenž realizuje posun signálu v dalším rozměru. Při výpočtu kumulantu 4. řádu přibývá ještě další proměnná, jenž realizuje posun ve čtvrtém rozměru.

5 OVĚŘENÍ VLASTNOSTÍ KUMULANTŮ

V předchozích studiích uvedených v kapitole 3 je uvedeno, že jsou kumulanty schopny snižovat variabilitu EKG cyklů patřících do jedné skupiny (klasifikační třídy) a tím usnadňovat analýzu EKG. Konkrétně se jednalo o potlačení rozdílů způsobených časovým posunem, rozdílnou výškou amplitudy cyklů a také potlačení šumu v signálech EKG. V následující části práce byly tyto situace namodelovány a z takto připravených signálů byly spočítány kumulanty druhého, třetího a čtvrtého řádu.

5.1 Vliv posunu po ose x (časová osa)

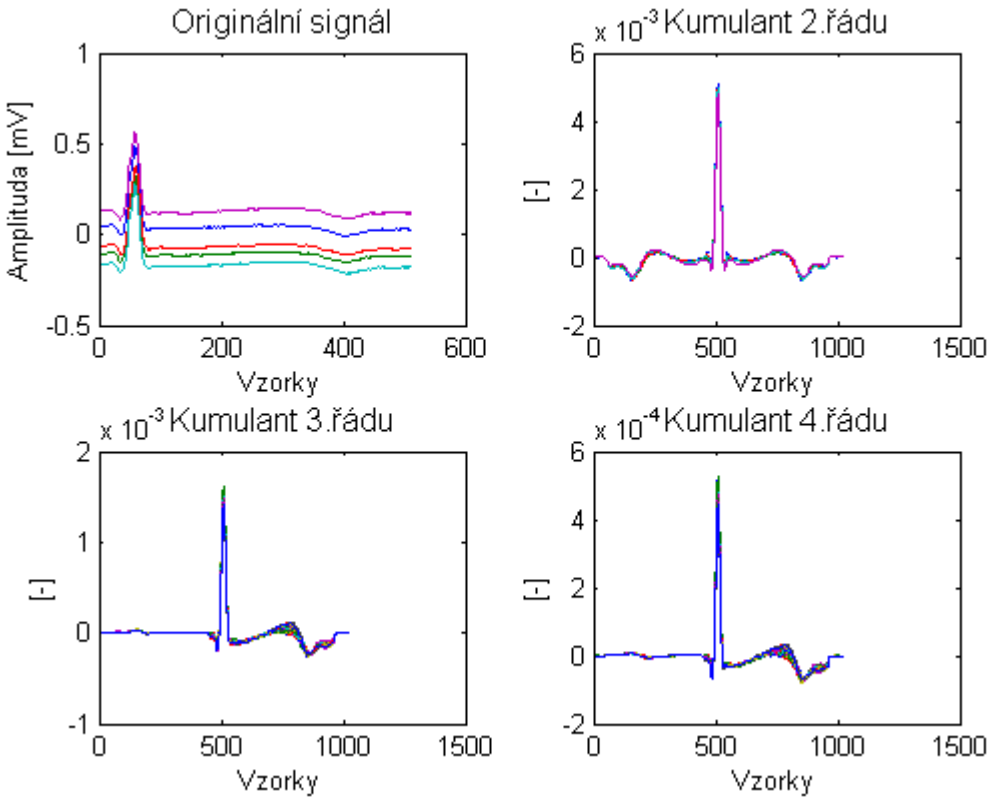
Na obr. 9 vidíme kumulanty, které byli vytvořeny ze signálu, kde byly EKG cykly *posunuty po časové ose x*, čehož bylo dosaženo přičtením několika nulových vzorků před začátek signálu a tomu odpovídajících nulových vzorků na konec signálu (kvůli zachování stejné délky). Poloha vlny R se pohybovala v rozmezí 80-220 vzorku signálu. Kumulanty všech řádů vykazují minimalizování variability jednotlivých cyklů, tedy v tomto případě minimalizování časového posunu. Nejlepší výsledek byl zaznamenán u kumulantů 3. řádu.



Obr. 9 Potlačení posunu po ose x (časový posun)

5.2 Vliv posunu po ose y (výchylka elektrogramu)

Na obr. 10 je zobrazeno vykreslení kumulantů vypočítaných ze signálu, kdy byli jednotlivé cykly vůči sobě *posunuty po ose y* nahoru i dolů v rozmezí 0.275 – 0.575 mV. Posun byl zajištěn přičtením náhodně vygenerované hodnoty pomocí funkce *randn* k celému cyklu. T Na první pohled vidíme, že rozdíl amplitudy kmitu R takto upravených cyklů byl značný. Již kumulanty druhého řádu velmi výrazně tyto rozdíly potlačili, jelikož vznikly cykly, jejichž rozdíl v amplitudě kmitu R byl zanedbatelný. Kumulanty třetího a čtvrtého řádu vykazují výsledky ještě lepší, mezi sebou navzájem srovnatelné. Dále si lze všimnout, že v originálních signálech velmi málo výrazné vlny T byly značně zvýrazněny (a to v tomto i předcházejícím případě), jejich použití by tedy zcela jistě usnadnilo například i detekci vlny T, při využití kumulantů pro hodnocení EKG záznamu.



Obr. 10 Potlačení posunu po ose y (výchylka elektrogramu)

5.3 Vliv přítomnosti Gaussovského šumu

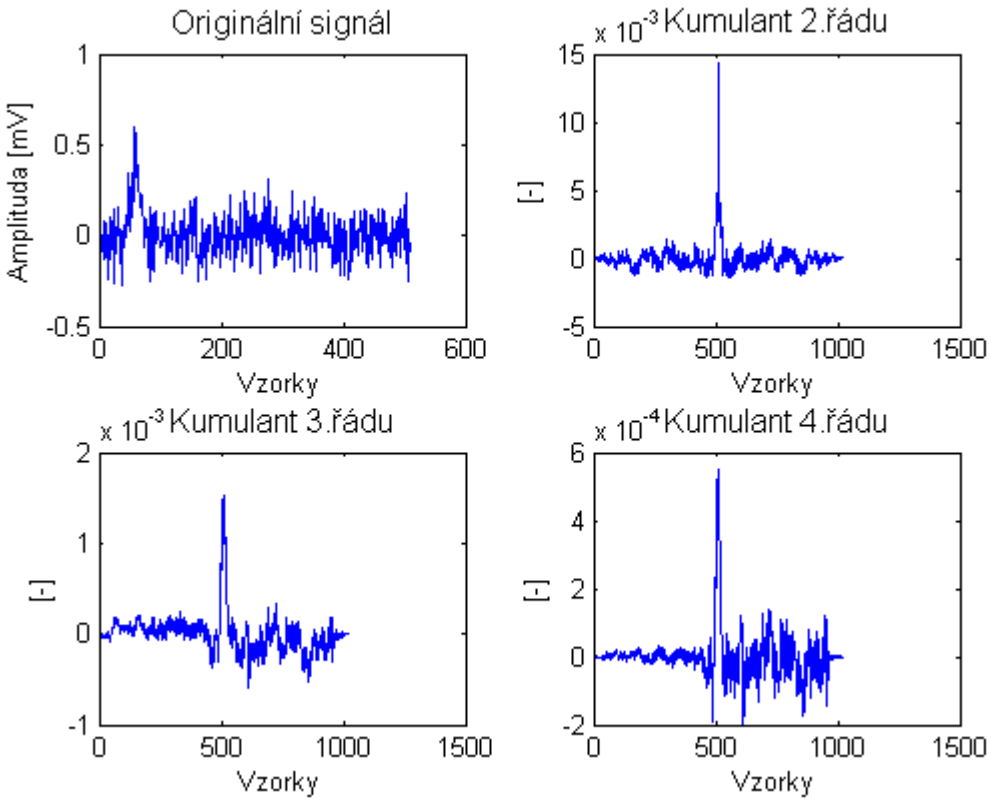
Obr. 11 potvrzuje, že vlastnosti kumulantů umožňují i *potlačení Gaussovského šumu.* Šum byl simulován náhodně generovaným bílým aditivním šumem s gaussovským rozložením, jehož parametry byly nastaveny tak, aby výsledná hodnota poměru signál/šum (angl.. signal-to-noise ratio, SNR) byla -3 dB. Takto vygenerovaný šum byl sečten

se vstupním signálem. Tím byla získána směs šumu a signálu, kde byl výkon signálu poloviční oproti výkonu šumu.

Výrazný šum v originálním signále, který by s největší pravděpodobností znesnadňoval jakékoliv další využití signálu, byl po vypočítání kumulantů značně potlačen. Výsledná hodnota SNR vypočítaná dle vzorce

$$SNR = 10 \cdot log \frac{\frac{1}{N}\sum_{n=0}^{N-1} s^2(n)}{\frac{1}{N}\sum_{n=0}^{N-1} s^2(n)}, \tag{16}$$

kde *s* značí signál a *n* šum je pro kumulant druhého řádu je 3,49 dB, což znamená, že výkon signálu byl více než dvojnásobný oproti výkonu šumu. U kumulantů řádu třetího a čtvrtého vyšla hodnota SNR ještě vyšší tj. 9,89 dB respektive 7,98 dB. Pokud se opět soustředíme na vlnu T, vidíme, že v původním signále bychom mohli jen těžko odhadovat, kde se nalézá. V případě kumulantů 2. řádu, je situace stejná, ovšem u kumulantů 3. řádu (respektive i 4. řádu) už informaci o umístění vlny T nacházíme.



Obr. 11 Potlačení šumu v signálu

Všechny tyto výše ukázané vlastnosti kumulantů byli potvrzeny i v předchozí studii [9].

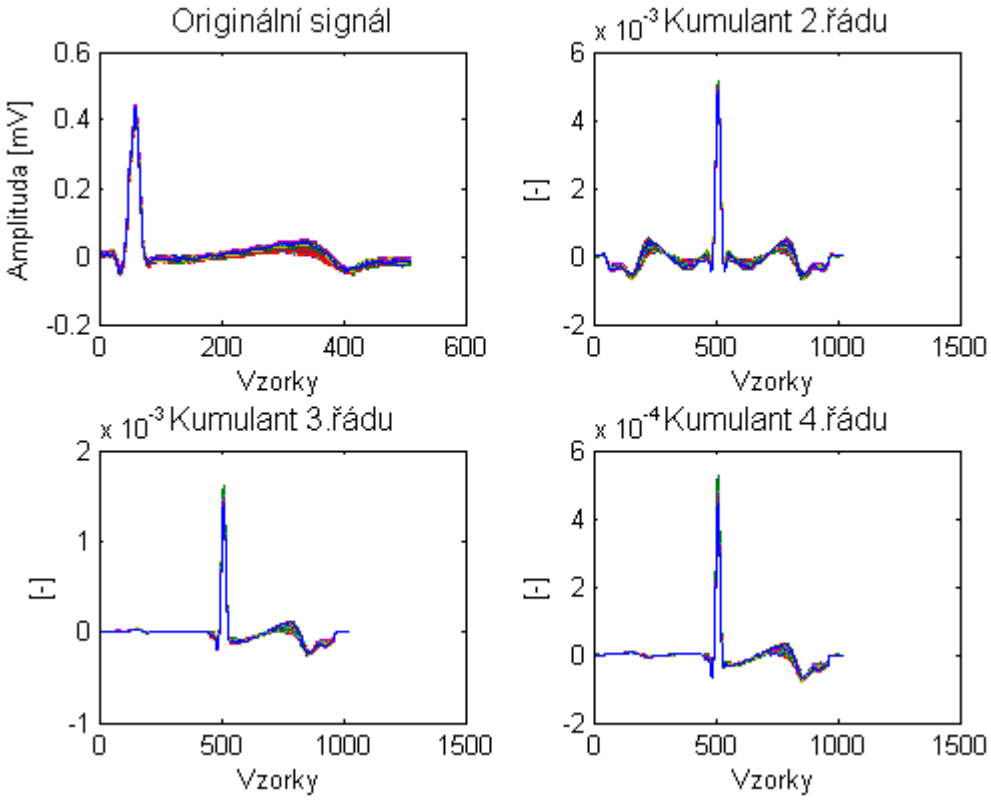
5.4 Výpočet kumulantů reálných dat

Na následujících obrázcích č. 12, 13, 14 a 15 jsou zobrazeny kumulanty 2., 3. a 4. řádu reálných EKG cyklů. Z roztřízených dat popsaných v kapitole 0 byly vybrány 4 skupiny reprezentativních cyklů, se kterými bylo dále pracováno. Skupiny cyklů byly pojmenovány následovně: Cykly Fyziologické, Ischemické, Ischemické 2 a skupina cyklů nazvaná Jiný typ (pravděpodobně se z velké většiny jedná o komorové extrasystoly). Skupiny Fyziologické, Ischemické a Ischemické 2 obsahovali 50 cyklů, skupina Jiný typ 61 cyklů. Tento nepříliš velký počet cyklů EKG byl zvolen z důvodu velké časové náročností výpočtu zejména kumulantů 3. a 4. řádu (výpočet kumulatnů 50ti cyklů 2. řádu trval průměrně pouze 1 sekundu, ovšem výpočet kumulantů 3. řádu 15 minut a kumulantů 4. řádu 25 minut). Pro naše potřeby je ale tento počet cyklů dostačující.

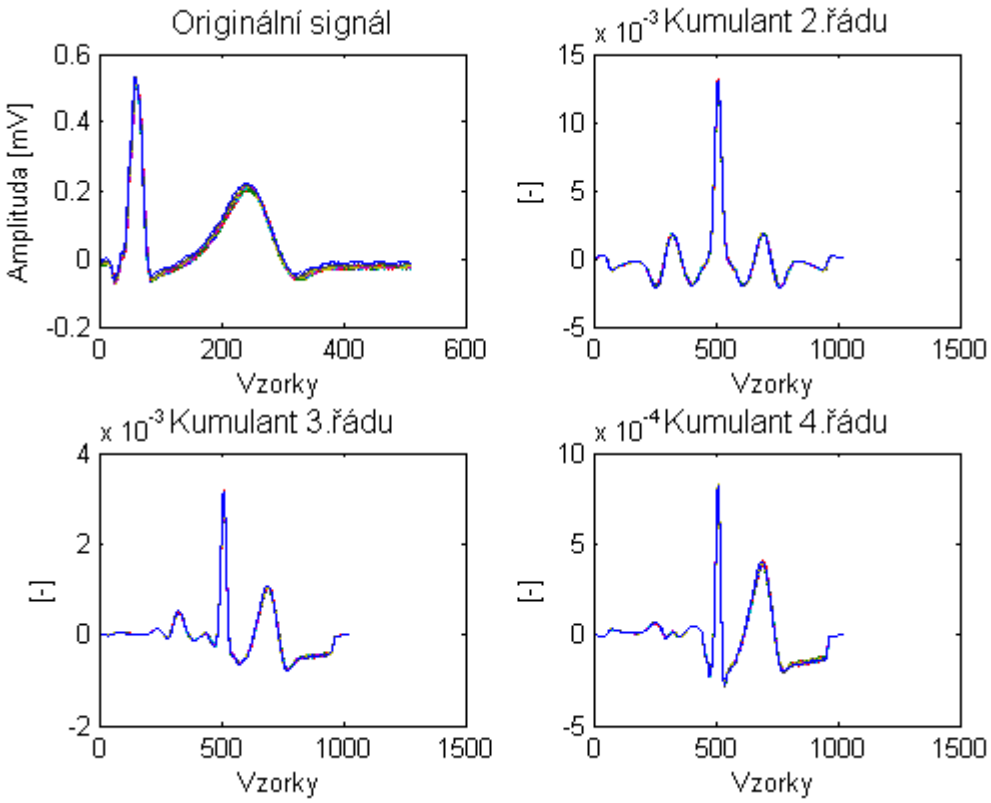
Nejvýraznější rozdíl ve variabilitě původního signálu oproti vypočítaným kumulatům lze již pouhým okem vidět na obrázku 14, kde byli kumulanty počítání ze skupiny cyklů Jiný typ s vysokou variabilitou. Poloha maximální amplitudy R vlny v originálním signálu byla v rozmezí od -0.1547 mV do 0.0678 mV, z čehož vyplývá variační šíře 0.2225 mV. V případě kumulantů byla variační šíře snížena na 0.0055 mV u kumulantu druhého řádu, v případě kumulantu třetího, respektive čtvrtého řádu na 0.013 mV a 0.0009 mV. Výrazně byl snížen i posun podél časové osy y, kdy v původním signálu vrchol kmitu R ležel v rozmezí 49 – 102 vzorku signálu, ale v korespondujících kumulantech byl vrchol amplitudy tohoto kmitu vždy na stejném vzorku.

Vzhledem k tomu, že variabilita byla již v rámci vybraných normálních (Obr.12) a ischemických skupin cyklů (Obr.13) malá, není zde snížení variability ihned natolik patrné jako v předcházejícím případě.

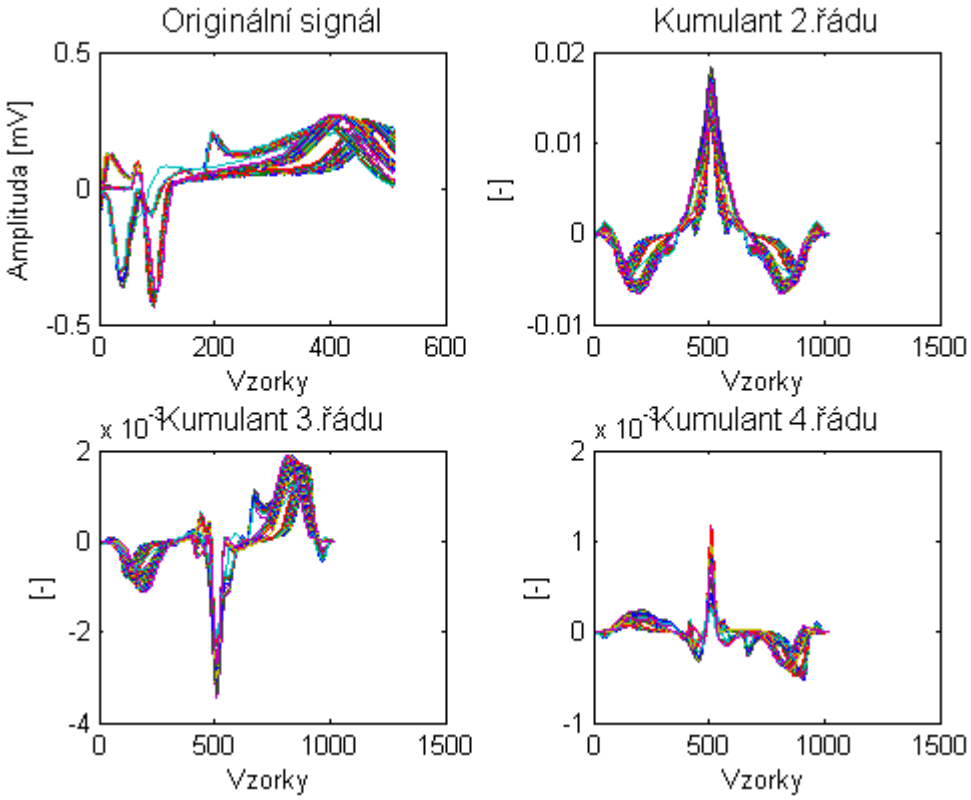
V kapitole č. 6 bude provedeno statistické zpracování získaných dat (korelační analýza). V kapitole č. 7 budou data využita k provedení dvou způsobů klasifikace cyklů do odpovídajících skupin pomocí umělé neuronové sítě.



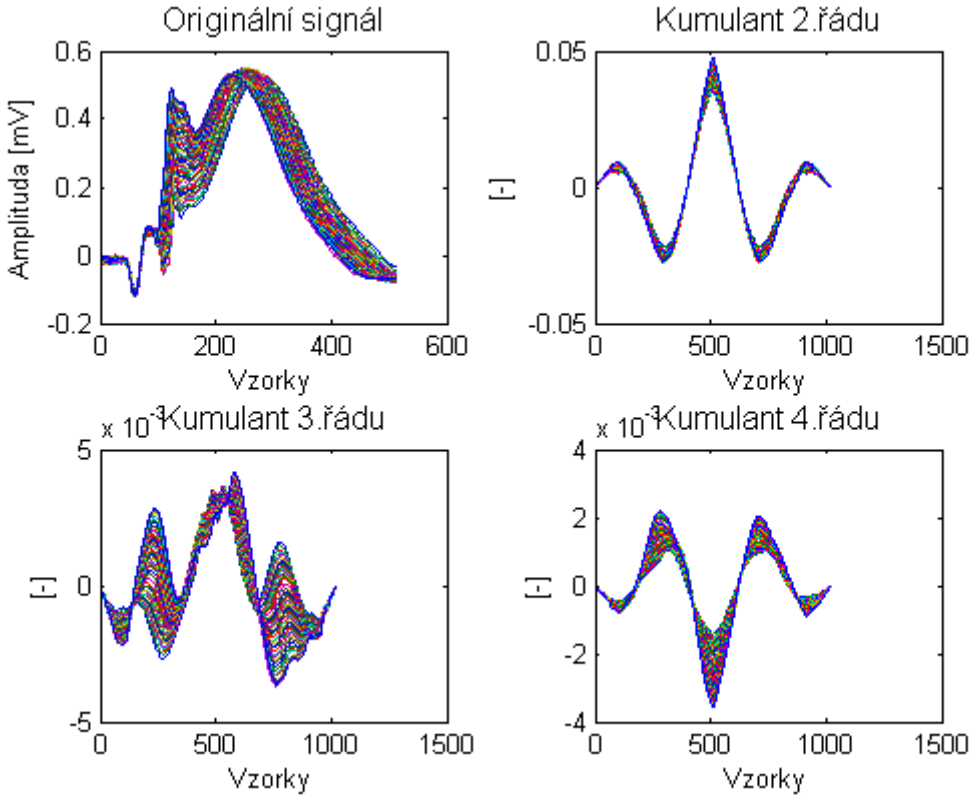
Obr. 12 Originální signál 50ti fyziologických EG cyklů a jejich kumulanty



Obr. 13 Originální signál 50ti ischemických EG cyklů a jejich kumulanty



Obr. 14 Originální signál 61 EG ektopických cyklů s vysokou variabilitou a jejich kumulanty(Jiný typ)



Obr. 15 Originální signál 50 EG silně ischemických cyklů a jejich kumulanty (Ischemické 2)

6 STATISTICKÉ ZHODNOCENÍ VARIABILITY DAT

Cílem této kapitoly je statistické potvrzení či vyvrácení výše zkoumaných vlastností kumulantů a vyhodnocení, který řád kumulantů vykazuje největší snížení variability a je tedy nejvhodnější ho použít pro následnou klasifikaci. Testování bude probíhat na sedmi skupinách různých typů cyklů (3 skupiny cyklů, na kterých probíhalo testování vlastností, a 4 skupiny dalších cyklů popsaných v předchozí kapitole).

6.1 Korelační analýza

Korelační analýza zkoumá míru závislosti a podobnosti dat. Konkrétně popisuje lineární vztahy mezi veličinami. Výstupem korelační analýzy je korelační koeficient r_{xy}, který se spočítá dle vzorce

$$r_{xy} = \frac{\sum x.y.\frac{1}{n}.\sum x.\sum y}{\sqrt{\left[\sum x^2.\frac{1}{n}.(\sum x^2)\right].\left[\sum y^2.\frac{1}{n}.(\sum y^2)\right]}}, \tag{17}$$

kde *x* a *y* značí porovnávaná data a *n* je délka vektorů dat. Hodnoty výsledného korelačního koeficientu se pohybují v intervalu <-1,1>, přičemž hodnota 1 značí přímou závislost (tedy v našem případě i podobnost signálu), hodnota -1 vyjadřuje zcela nepřímou lineární závislost – antikorelaci. Hodnota korelačního koeficientu rovna nule znamená, že mezi procesy není žádná lineární závislost [2].

Pomocí funkce *corcoeff* byly spočítány korelační koeficienty všech 7 typů cyklů a jejich kumulantů. Nejprve byla získána matice korelačních koeficientů. Z této matice hodnot byla spočítána střední hodnota a směrodatná odchylka korelačního koeficientu pro všechny skupiny cyklů. Výsledky jsou uvedeny v tabulce 1 v další kapitole. Pro každý typ cyklů je v tabulce zvýrazněno, který kumulant vykazoval nejvyšší korelaci mezi cykly.

6.1.1 Výsledky korelační analýzy

Z korelační analýzy vyplývá, že u většiny zkoumaných skupin cyklů došlo u všech řádů kumulantů ke zvýšení hodnot korelace oproti původnímu signálu a tedy i ke snížení variability ve skupině. Pouze u skupiny Jiný typ toto nebylo zcela potvrzeno, jelikož hodnota korelace byla v případě 3. a 4. řádu kumulantu nižší než u původního signálu. Ve dvou případech byla největší korelace u kumulantů 2. řádu, kumulant 3. řádu vykazoval nejvyšší korelaci ve třech případech, stejně jako kumulant řádu 4, přičemž v jednom případě byli hodnoty korelace kumulantů řádů 3 a 4 zcela totožné. Výsledky korelační analýzy jsou uvedeny v tabulce 1 a nejvyšší hodnoty korelace pro každý typ cyklů jsou šedě zvýrazněny.

Tabulka 1 Korelační koeficienty a jejich směrodatné odchylky původních skupin cyklů a korespondujících kumulantů

Řád kumulantu / Cykly	Původní signál	Kumulant 2. řádu	Kumulant 3. řádu	Kumulant 4. řádu
Posunuté po ose X	0.1938 ± 0.0018	0.9956 ± 0.0035	0.9976 ± 0.0018	0.9976 ± 0.0018
Posunuté po ose Z	0.9977 ± 0.2797	0.9974 ± 0.0025	0.9990 ± 0.0008	0.9990 ± 0.0007
Zašuměné	0.4539 ± 0.2678	0.7954 ± 0.1098	0.8585 ± 0.0748	0.7830 ± 0.1133
Fyziologické	0.9927 ± 0.0040	0.9951 ± 0.0078	0.9956 ± 0.0048	0.9957 ± 0.0045
Ischemické	0.9987 ± 0.0009	0.9995 ± 0.0004	0.9995 ± 0.0003	0.9994 ± 0.0004
Jiný typ	0.7337 ± 0.2694	0.9422 ± 0.0584	0.8682 ± 0.1036	0.8268 ± 0.1249
Ischemické 2	0.9422 ± 0.0691	0.9988 ± 0.0014	0.8431 ± 0.1754	0.9942 ± 0.0059

Na základě korelační analýzy bylo vzhledem k vysokým hodnotám korelace v případě všech řádů kumulantu a zároveň malým rozdílům mezi těmito hodnotami určeno, že pro případnou klasifikaci je dostačující využít kumulant co nejnižšího, tedy druhého, řádu a to z důvodu nejméně náročného výpočtu.

Jinou, složitější, ale o to přesnější, variantou by byla možnost klasifikace na základě několika nebo všech kumulantů zároveň (v případě zařazení různých kumulantů stejného cyklu do různých skupin, je konečné zařazení cyklu do odpovídající skupiny provedeno podle toho řádu kumulantu, který v předchozím zkoumání vykazoval nejlepší výsledky). Tato možnost byla využita například v [9].

6.2 Příznaky odvozené z kumulantů

V této kapitole budou z původních cyklů a jejich kumulantů odvozeny statistické parametry (tzv. příznaky), které budou dále sloužit pro reprezentaci cyklů při jejich následné klasifikaci. Zvoleny byly tři parametry: variační koeficient, koeficient šikmosti a parametr vyjadřující normalizovanou sumaci. Obecný význam a způsob výpočtu vybraných parametrů je popsán níže (viz. také [27]).

1. Koeficient symetrie (SYM) - je statistický parametr, který charakterizuje pravděpodobnostní rozdělení dat. Dané rozdělení porovnává s normálním rozdělením pravděpodobnosti. Je definován jako

$$SYM = \frac{\sum_{l=1}^{n} x(l) - x(n-l)}{\sum_{l=1}^{n} x(l)}, \tag{18}$$

kde n je délka vektoru hodnot cyklu *x*.

2. Variační koeficient (VK) - je charakteristikou variability náhodné veličiny. Udává tedy míru variabilitu jednotlivých EKG cyklů v rámci jedné klasifikační skupiny. Jeho hodnotu získáme výpočtem dle

$$VK = \frac{\sqrt{D(x)}}{|E(x)|}, \tag{19}$$

kde *x* je vektor hodnot jednoho cyklu, *D(x)* směrodatná odchylka a *E(x)* střední hodnota.

3. Normalizovaná sumace (NS) - výpočet normalizované sumace je

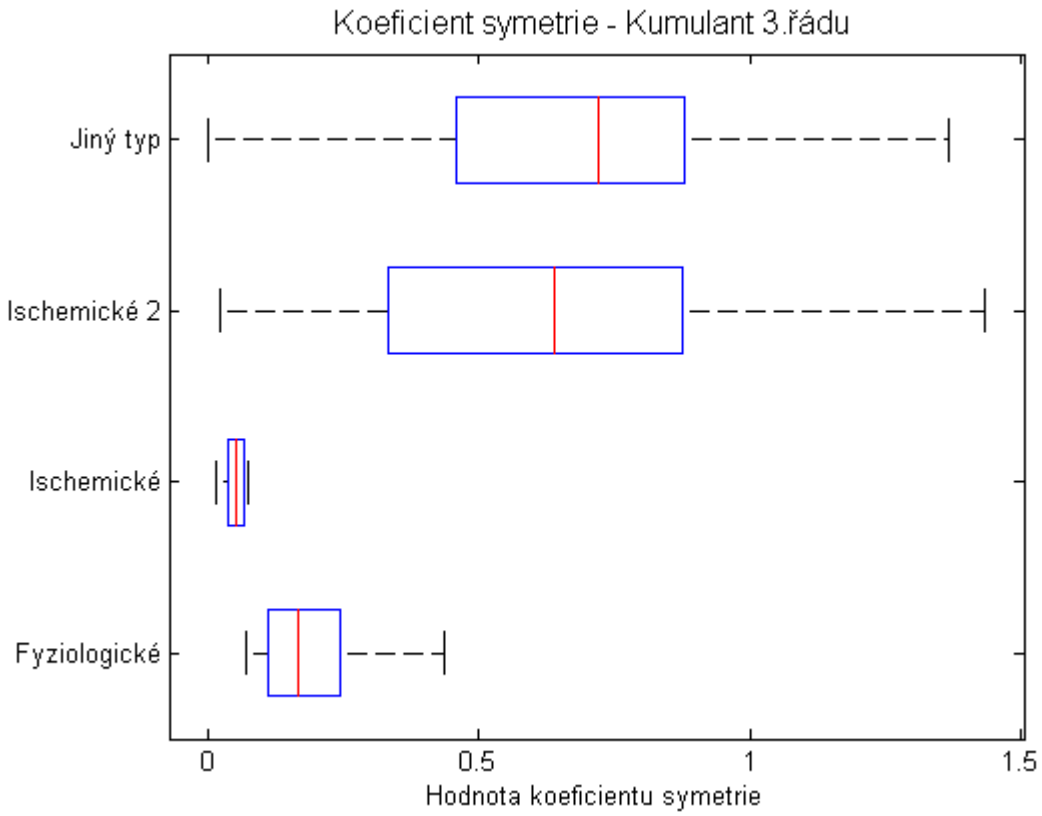
$$NS = \frac{\sum x}{\sum |x|}, \tag{20}$$

kde x je vektor hodnot jednoho cyklu.

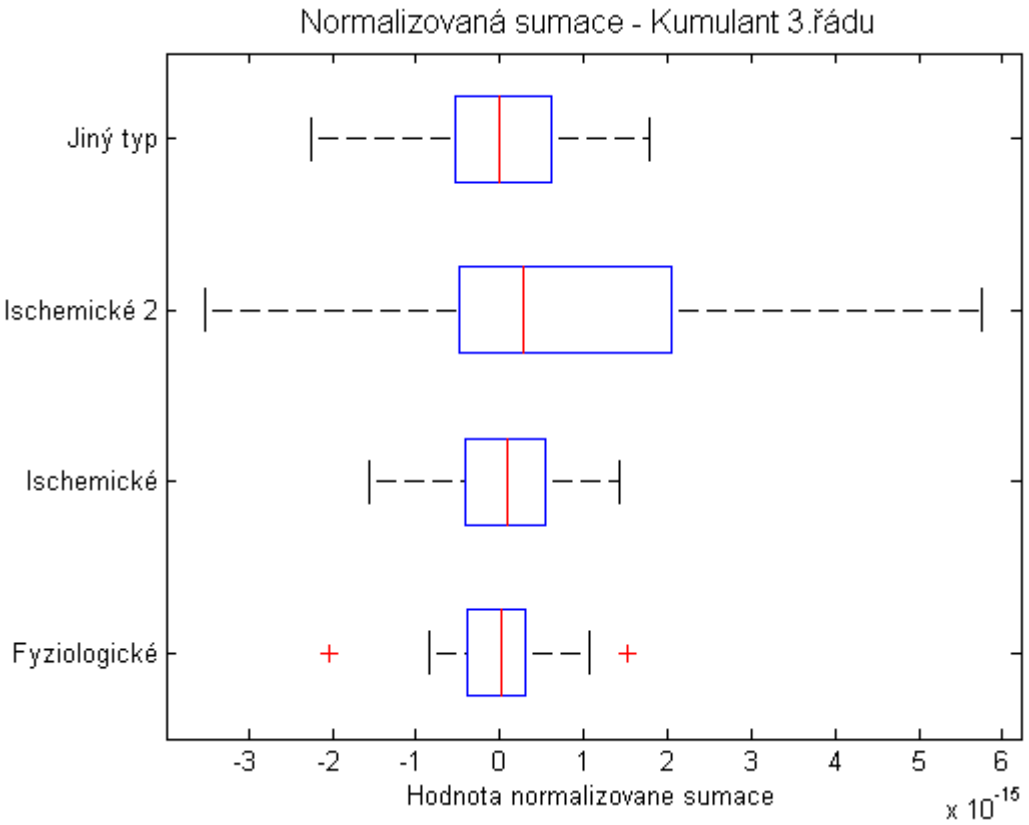
Na základě vzorců č. 18, 19, 20 byly provedeny výpočty uvedených parametrů pro původní cykly i všechny řády kumulantů. Takto byly pro každý cyklus získány tři hodnoty, které budou sloužit jako vstupní data pro klasifikátor. Před použitím hodnot parametrů pro klasifikaci cyklů bylo třeba ověřit, zda jsou hodnoty parametrů v rámci různých skupin cyklů odlišné a bude možné na jejich základě klasifikaci vůbec provést.

K tomuto ověření bylo využito krabicových grafů, jejichž vykreslení bylo realizováno pomocí funkce *boxplot*. Krabicové grafy graficky vyjadřují rozdělení hodnot a jejich variability. Přímka uvnitř boxu představuje medián. Horní hrana obdélníku odpovídá hornímu kvartilu, dolní hrana kvartilu dolnímu. Mezi horním a dolním kvartilem leží 50% hodnot a tato oblast se označuje jako interkvartilový interval (IQR). Koncové úsečky představují extrémní hodnoty (1,5 × IQR). Hodnoty, které se nacházejí ve větší vzdálenosti než 1,5 × IQR od mediánu jsou reprezentovány červenými křížky. [20]

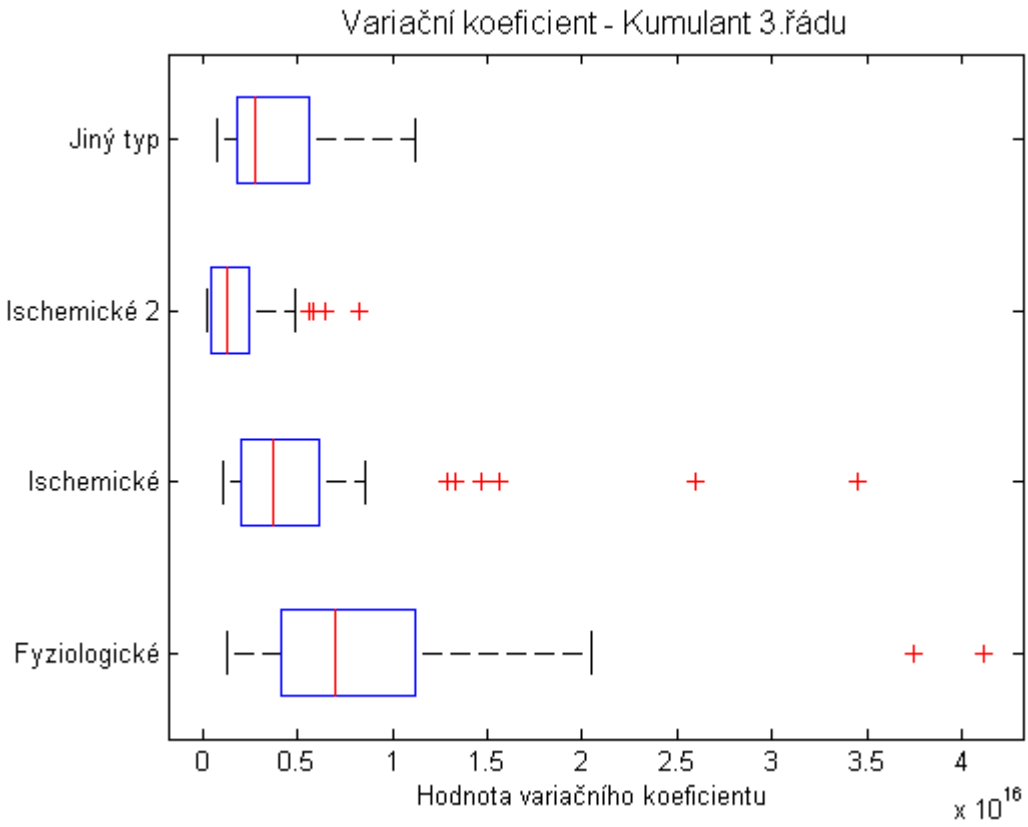
Na základě těchto grafů a zejména rozložení hodnot patřících do intervalu IQR bylo vyhodnoceno, že v případě koeficientu symetrie jsou hodnoty parametrů mezi různými typy cyklů dobře rozlišitelné (hodnoty parametrů leží v jasně odlišných rozmezí). V případě hodnot normalizované sumace a variačního koeficientu je rozlišitelnost menší (hodnoty různých skupin cyklů se překrývají), z toho důvodů není jisté, zda bude klasifikace správně fungovat. V případě variačního koeficientu je možno si všimnout dalšího problému, a to četného výskytu hodnot parametrů ležícími ještě za extrémními hodnotami, které by mohli negativně ovlivňovat klasifikace. Stejné závěry o rozlišitelnosti lze vyvodit i pouhým porovnáním konkrétních hodnot parametrů. Níže jsou uvedeny příklady hodnocených krabicových grafů pro kumulanty 3. řádu.



Obr. 16 Krabicové grafy koeficientů symetrie pro kumulant 3. řádu



Obr. 17 Krabicové grafy normalizované sumace pro kumulant 3. řádu



Obr. 18 Krabicové grafy variačních koeficientů pro kumulant 3. řádu

7 KLASIFIKACE DAT POMOCÍ UMĚLÉ NEURONOVÉ SÍTĚ

V této kapitole budou provedeny a zhodnoceny dvě metody klasifikace EKG cyklů do odpovídajících klasifikačních skupin pomocí umělé neuronové sítě. V prvním případě se bude jednat o v praxi běžně využívaný způsob klasifikace EKG cyklů, kdy je vstupem pro klasifikátor celý průběh cyklu (kumulantu) případně několik bodů daného signálu rozložených ekvidistantně po jeho délce (signál je podvzorkován). Tohoto způsobu klasifikace bylo využito např. v [9], [16], [17]).

Jinou možností klasifikace EKG cyklů je klasifikace na základě odvození několika statistických parametrů (příznaků) z původních cyklů a jejich kumulatnů. Jako vstup pro klasifikátor je pak použit vektor hodnot parametrů reprezentujících cykly, místo celého průběhu kumulantu (původního cyklu). Tento způsob klasifikace je výhodnější z toho důvodu, že je pracováno s mnohem menším počtem dat a klasifikace je tedy technicky méně náročná a rychlejší. V dřívějších studiích bylo tohoto principu využito v [27].

Data, která budou klasifikována, jsou popsána a zobrazena v kapitole 4.4. Vstupními daty tedy bude 211 cyklů čtyř různých typů.

7.1 Neuronová síť

7.1.1 Obecný popis neuronových sítí

Umělá neuronová síť (angl. artifical neural network, ANN) vznikla na podkladě výzkumu lidského mozku, je tedy umělou analogií k biologické neuronové síti. Neuronové sítě patří mezi typicky používané klasifikační algoritmy. Informace pro tuto část práce byly čerpány z [3] a [13].

Základním výkonným prvkem neuronové sítě je neuron. Do neuronu vstupuje velký počet vstupů, výstup je vždy pouze jeden (může se ale dále větvit a připojovat na více dalších neuronů). Neuron lze obecně popsat dle vztahu

$$y = f[\sum_{i=0}^{N} w_i x_i - \theta], \tag{18}$$

kde x je vstupní vektor hodnot, w vektor vah, f aktivační funkce neuronu a θ je práh neuronu. Váhy w_i u každého neuronu představují jeho lokální paměť. Spojením všech neuronů dostáváme celkovou paměť sítě. Aktivační funkce převádí vnitřní potenciál neuronu (výraz v závorce v rovnici 18) do definovaného oboru výstupních hodnot.

Umělá neuronová síť pracuje ve dvou fázích. První fází je fáze adaptivní, kdy se síť učí. Dochází k tomu úpravou vah. Druhá je fáze aktivní, kdy dochází k vybavování naučených informací a díky tomu ke klasifikaci cyklů.

V této práci byla použita dopředná neuronová síť se zpětným šířením chyby (viz např. [13]). Učení sítě tohoto typu sítě probíhá následujícím způsobem. Síti předložíme učební (trénovací) množinu vstupů a jim odpovídajících výstupů. Váhy nastavíme na nulové, případně jiné, náhodné, hodnoty. Po proběhnutí první epochy získáme chybu výstupní vrstvy, která se zpětně šíří (přepočítává) do předchozích vrstev a podle její hodnoty se upravují váhy neuronů. Tímto způsobem dochází k minimalizování odchylky mezi skutečnými a požadovanými výstupy použité trénovací množiny. Tímto způsobem pokračuje učení sítě až do dosažení maximálního počtu epoch, nebo do dosažení minimální přípustné hodnoty chyby sítě. Výstupem ANN je zařazení EKG cyklů do odpovídající skupiny.

7.1.2 Použitá neuronová síť

Cykly EKG byly před použitím ANN rozděleny do tři množin. První množinou je množina trénovací (obsahuje data, na kterých se síť učí), validační (data sloužící pro průběžné testování úspěšnosti klasifikace a případnému předčasnému ukončení učení na základě dosažení přípustné chyby sítě) a množina testovací. Hodnocení testování úspěšnosti klasifikace bude probíhat na jak na trénovací množině dat, která obsahuje 33% dat z celkového počtu vybraných dat (jedná se tedy o stejná data, byla použita pro učení sítě) a na testovací množině dat, která obsahuje 50% dat. Výstup sítě při klasifikaci cyklů do 4 skupin je kódován jako 1000 v případě zařazení cyklu do první klasifikační skupiny, jako 0100 v případě zařazení cyklu do druhé klasifikační skupiny atd.

Hodnocena bude tedy úspěšnost klasifikace, která odpovídá procentuálnímu vyjádření počtu správně zařazených cyklů do odpovídajících klasifikačních skupin.

Nejdůležitější parametry použité ANN jsou:

- typ: vícevrstvá dopředná neuronová síť se zpětným šířením chyby;
- topologie: 3 vrstvy, v první (distribuční) vrstvě je tolik neuronů, kolik hodnot obsahuje vstupní vektor (tzn. původní cykly 510, kumulanty 1019, parametry 3 neurony), v druhé vrstvě (tedy první skryté vrstvě) bylo použito 5 neuronů, třetí (tzn. výstupní) výstupní vrstva obsahuje tolik neuronů, do kolika klasifikačních skupin je třízeno (tj. v případě klasifikace na základě celých průběhů cyklů 4 neurony, v případě klasifikování dle parametrů 2 neurony);
- aktivační funkce neuronu: v distribuční vrstvě: sigmoidální (jedná se o spojitou monotónní funkci se spojitými derivacemi a jednoduchým způsobem výpočtu gradientu); v skryté vrstvě: lineární (zvolena z toho důvodu, že požadovaný výstup není pouze v intervalu od -1 do 1);

- algoritmus pro učení: minimalizace chyby metodou gradientního sestupu chybové funkce;
- počet epoch, po kterém je zastaveno učení: 1000,
- přípustná chyba sítě: 0,001,
- iniciační hodnoty vah: nulové.

7.2 Klasifikace na základě celých průběhů

V tabulce č. 2 jsou uvedeny výsledky klasifikace, kdy byly na vstup klasifikátoru přivedeny celé průběhy hodnocených signálů. Vstupní data byla třízena do 4 klasifikačních skupin**.** Úspěšnost klasifikace na základě původních průběhů cyklů byla při klasifikaci cyklů pocházejících z trénovací množiny 74,3 %, v případě klasifikace cyklů z testovací množiny dat 72,9 %. V případě kumulantů 3. řádu byla úspěšnost výrazně zvýšena a to až na hodnoty 97,1 % a 95,1 %. Zhruba o 10% nižší úspěšnost byla při klasifikaci cyklů podle kumulantů 4. řádu. Výsledky klasifikace kumulantů řádu 2. byli o několik málo procent horší než u původních průběhů cyklů.

Tabulka 2 Výsledky klasifikace na základě celých průběhů

	Úspěšnost klasifikace [%]	
	trénovací množina	**testovací množina**
Původní data	74,3	72,9
Kumulant 2. řádu	70,1	68,4
Kumulant 3. řádu	97,1	95,1
Kumulant 4. řádu	85,7	85,7

Tyto výsledky je možno porovnat s výsledky získaným v [8]. Úspěšnost klasifikace v této studii byla 4,2 %, 54,2 % a 20,8 % v pořadí kumulant 2., 3. a 4. řádu. Výrazně lepší výsledky dosažené zde, se dají vysvětlit tím, že ve studii [8] bylo použito mnohem větší skupiny dat (9367 cyklů) na základě čehož lze předpokládat i větší variabilitu původních dat a následně i kumulantů a z toho vyplývající obtížnější klasifikaci. Výsledky obou studií se shodují v tom, že nejvyšší úspěšnost klasifikace vykazuje kumulant 3. řádu a nejnižší kumulant 2. řádu.

7.3 Klasifikace na základě příznaků odvozených z kumulantů

Dalším provedeným způsobem klasifikace EKG cyklů je klasifikace založená na vektoru odvozených parametrů reprezentujících cykly. Popis použitých odvozených parametrů (variační koeficient, koeficient symetrie, normalizovaná sumace) je uveden v kapitole 6.2.

Nejprve byla vyzkoušena klasifikace všech vstupních dat najednou (tedy třízení 211 cyklů reprezentovaných třemi parametry do 4 klasifikačních skupin). Úspěšnost takovéto klasifikace se ovšem pohybovala kolem 20-30% což je zcela nedostačující. První příčinou byl zřejmě nesprávný výběr parametrů, kterými byla data reprezentována, o čemž vypovídají i výsledky testování odlišnosti hodnot parametrů v kapitole 6.2.1., které nebyli dostačující. Za další důvod objasňující takto nízkou klasifikační úspěšnost, je zřejmě možno považovat zcela nedostatečný počet testovaných dat. Obdobný způsob využívající klasifikace na základě parametrů byl proveden ve studii [27]. Zde bylo dosaženo klasifikační úspěšnosti kolem 98 %. Takto vysoké úspěšnosti bylo dosaženo z toho důvodu, že bylo použito 30 různých parametrů, které reprezentovali jednotlivé cykly. Výpočet tak velkého množství parametrů by ale nebyl v rozsahu této práce reálný.

Z výše uvedeného důvodu byly jako vstupní data pro ANN vždy postupně použity pouze dva typy cyklů ve všech možných kombinacích a klasifikováno bylo tedy pouze do dvou skupin. V tabulce č. 3 jsou uvedeny výsledky při použití tohoto způsobu klasifikace. Klasifikační úspěšnost různých kombinací cyklů v rámci stejných typů dat se vždy mírně lišila, proto byla spočítána průměrná hodnota, aby bylo možné výsledky snadněji zhodnotit.

Tabulka 3 Výsledky klasifikace na základě odvozených parametrů

	Úspěšnost klasifikace [%]							
	Původní data		**Kumulant 2.ř**		**Kumulant 3.ř.**		**Kumulant 4.ř.**	
	tren.	**test.**	**tren.**	**test.**	**tren.**	**test.**	**tren.**	**test.**
Fyz. x isch.	93,9	90,9	87,9	84,4	90,9	87,8	87,9	84,8
Fyz. x isch2.	86,1	83,8	75,1	73,2	86,1	83,8	86,5	83,3
Fyz. x jiné	100	100	78,8	60,6	81,8	81,8	72,7	69,7
Isch. x isch2.	97,3	100	94,1	83,4	89,2	88,9	89,2	83,3
Isch. x jiné	100	100	78,7	72,6	90,9	87,9	81,8	81,8
Isch2. x jiné	100	100	72,4	71,3	76,5	74,1	62,2	61,1
Průměr	**96,2**	**95,8**	**81,2**	**74,3**	**85,9**	**84,1**	**80,1**	**77,3**

Tren.: trénovací množina dat, test.: testovací množina, fyz.: fyziologické cykly, isch.: ischemické cykly, isch2.: ischemické cykly 2.

Nejvyšší úspěšnosti klasifikace bylo dosaženo ve všech případech při využití vektoru parametrů, jenž reprezentoval původní cykly. K tomuto došlo pravděpodobně z toho důvodu, že hodnoty původních dat se pohybují v řádech jednotek, ovšem hodnoty kumulantů v mnohem nižších řádech (10^{-4} – 10^{-9}), což je zřejmě jedním z důvodů, který ovlivnil nižší rozlišitelnost hodnot parametrů z kumulatnů vypočtených. Z tohoto důvodu nebudeme na základě tohoto způsobu klasifikace hodnotit úspěšnost klasifikace původních cyklů oproti kumulantů, ale pouze porovnáme úspěšnost klasifikace pouze v rámci kumulantů. Opět lze vidět, že nejvyšší úspěšnosti v rámci různých řádů kumulantů bylo stejně jako v případě klasifikace na základě celých průběhů cyklů dosaženo v případě kumulantu 3. řádu. Výsledky dosažené v případě kumulantů 2. a 4. řádu jsou navzájem srovnatelné.

ZÁVĚR

V této práci byla zkoumána využitelnost kumulantů vyšších řádů pro analýzu EKG cyklů. K výzkumu byly využity signály získané v průběhu experimentu [11], který byl prováděn na izolovaných králičích srdcích a docházelo při něm ke střídání fází ischemie a reperfuze. Pro výpočet kumulantů bylo využito programovací rozhraní Matlab.

Prvním cílem této práce bylo seznámení se s vlastnostmi a způsoby výpočtu kumulantu druhého až čtvrtého řádu se zaměřením na jejich využitelnost pro analýzu EKG. Tato problematika a popis základních vlastností kumulantů je rozebrána v kapitole 2. Následující kapitola se věnuje literární rešerši dané problematiky – je zde shrnuto využití kumulantů v současné praxi. Realizace výpočtu kumulantů a popis použitých funkcí jsou uvedeny v kapitole 4.

Dalším cílem práce bylo ověření předpokládaných vlastností kumulatnů výhodných pro analýzu EKG signálu (viz kap. 5). Na modelovaných situacích bylo ověřeno, že výpočtem kumulantů z EKG cyklů lze potlačit rozdílný časový posunu jednotlivých cyklů, který může být způsoben například nepřesnou detekcí maxima R vlny. Dále došlo také k snížení různého posunu cyklů podél osy y (tj. potlačení kolísání nulové izolinie) vzniklého například „plaváním elektrod“, ke kterému dochází kvůli nepevně uchyceným elektrodám. Kumulanty také dokázaly částečně potlačit Gaussovský šum, kdy došlo k snížení SNR z hodnoty -3 dB na hodnotu 3,49 dB respektive 9,89 dB a 7,98 dB u odpovídajících kumulantů. Potvrzené vlastnosti kumulantů lze dobře využít ke snížení variability EKG cyklů v rámci jedné klasifikační skupiny, z čehož vyplývá možnost snadnější a přesnější analýzy EKG, následného vyhodnocení záznamu a stanovení diagnózy.

V rámci práce byla také vytvořena databáze EKG cyklů roztřízených podle morfologických znaků, mezi které patřil například sklon úseku ST a tvar vlny T. Nejprve byly cykly tříděny v rámci jednotlivých fází experimentu (obr. 4-6) a poté byly tyto určité typy cyklů z důvodu velké podobnosti mezi některými typy komplexně sloučeny do výsledných šesti typů, přičemž čtyři z nich patří do klasifikační skupiny „normální cykly“ a dva do klasifikační skupiny „ischemie“ (viz obr. 7). Bližší popis třídění spolu s vykreslením všech typů cyklů je uveden v kapitole 0. Z těchto dat byly vybrány 4 reprezentativní skupiny cyklů (Fyziologické, Ischemické, Ischemické 2 a Jiný typ), z kterých byly spočítány kumulanty 2., 3. a 4. řádu (viz obr. 12-15). Všechny skupiny obsahovali přibližně 50 cyklů. Tento nepříliš velký počet cyklů byl zvolen z důvodu velké časové náročnosti výpočtu. Velká časová náročnost výpočtu se dá považovat za fakt, který by mohl snižovat perspektivu použití kumulantů v praxi V další části práce bylo těchto 4 vybraných skupin cyklů využito pro další testování.

V další části práce (kapitola 6.1) byla provedena korelační analýza. Bylo zjištěno, že ke zvýšení hodnot korelace oproti původním cyklům dochází u všech řádů kumulantů. Vzhledem k tomu, že hodnoty všech korelačních koeficientů všech kumulantů byly vysoké, bylo na základě korelační analýzy stanoveno, že by mělo být dostačující pro klasifikace EKG cyklů využit kumulant co nejnižšího řádu, a to vzhledem k menší náročnosti výpočtu. V rámci kapitoly 6.2 byly spočítány statistické parametry pro reprezentaci cyklů, které byly dále využity v kapitole 7.3 pro klasifikaci.

V poslední části práce byly pomocí umělé neuronové sítě provedeny dva způsoby klasifikace EKG cyklů. V prvním případě byla klasifikace založena na celých průbězích kumulantů (kap. 7.2) a v druhém případě na byla odvozených statistických parametrech (kap. 7.3). V kapitole 7.1 je uveden základní popis neuronových sítí a také jsou zde uvedeny parametry použité neuronové sítě. Z obou testovaných klasifikačních metod vyplývá, že při klasifikaci EKG cyklů podle kumulantů řádů 3. a 4. dochází k výraznému zvýšení klasifikační úspěšnosti (tj. úspěšnějšímu zařazení cyklů do odpovídajících skupin). V případě využití kumulantu 2. řádu ke zvýšení úspěšnosti klasifikace ve většině případů nedochází. Výsledky obou způsobů klasifikace byly porovnány s předchozími studiemi, které byly provedeny srovnatelnými způsoby a za obdobných podmínek.

V návaznosti na tuto práci by bylo možné provést výpočet kumulantů pro větší počet cyklů a více různých klasifikačních skupin, data podrobit klasifikaci a tyto výsledky zhodnotit. Zajímavé by bylo vyzkoušet klasifikaci na základě kombinace různých řádů kumulantů a mimo jiné porovnat technickou náročnost tohoto způsobu vůči zpřesnění klasifikace. V případě klasifikace na základě parametrů odvozených z kumulantů by bylo vhodné použít větší počet různých parametrů. Dále by bylo možno detailněji diskutovat velkou časovou náročnost výpočtu kumulantů 3. a 4. řádu vzhledem k reálnému použití.

Závěrem lze říci, že kumulanty vyšších řádu mají v oblasti analýzy EKG velkou perspektivu a bylo by užitečné je v reálné praxi pro klasifikaci využít.

SEZNAM POUŽITÉ LITERATURY

[1] AGRAFIOTI, F. a D. HATZINAKOS. ECG Based Recognition Using Second Order Statistics. In: ILOW, J. *Proceedings: CNSR 2008 Sixth Annual Conference on Communication Networks and Services Research [online]*. Los Alamitos, Calif., 2008 [cit. 2012-11-02]. ISBN 978-0-7695-3135-9. Dostupné z: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4519843

[2] ANDĚL, J. *Matematická statistika*. 2. vyd. Praha: SNTL, 1985, 346 s.

[3] BEALE, M. H., M. T. HAGAN a H. B. DEMUTH. *Neural Network Toolbox*: *User´s Guide* [online]. [cit. 2012-12-07]. Dostupné z: http://www.mathworks.com/help/pdf_doc/nnet/nnet_ug.pdf

[4] HABERL, R. *EKG do kapsy.* 1. české vyd. Praha: Grada, 2012, 281 s. ISBN 9788024741925

[5] HAMAN, P. *Výukový web EKG. Základy EKG [online]*. [cit. 2012-11-09]. Dostupné z: http://ekg.kvalitne.cz/start.htm.

[6] HAMPTON, J. R. *EKG v praxi.* 2. české vyd. Překlad Eliška Potluková. Praha: Grada, 2007, 362 s. ISBN 978-80-247-1448-6.

[7] HAMPTON, J. R. *EKG stručně, jasně, přehledně.* 1. vyd. Překlad Jiří Dostal, Barbora Kadlecová. Praha: Grada, 1996, 107 s. ISBN 80-716-9153-4.

[8] HANNESSCHLAGER, G. *HOSA – Higher Order Spectra Analysis.* 2004. [cit. 2012-10-15]. Dostupné z: http://www.emt.jku.at/education/Inhalte/se_moderne_methoden/WS0405/hannesschlaeger_hosa.pdf. 2004

[9] KARIMIFARD, S. a A. AHMADIAN. *A robust method for diagnosis of morphological arrhythmias based on Hermitian model of higher-order statistics [online]*. [cit. 2012-10-15]. DOI: 10.1186/1475-925X-10-22. Dostupné z: http://www.biomedical-engineering-online.com/content/10/1/22

[10] KHADRA, L., A. S. AL-FAHOUM a S. BINAJJAJ. A Quantitative Analysis Approach for Cardiac Arrhythmia Classification Using Higher Order Spectral Techniques. *IEEE transactions on bio-medical engineering / Bio-medical Engineering Group [online]*. 2005, no. 11, s. 1840-1845 [cit. 2012-10-27]. ISSN 0018-9294. Dostupné z: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1519592

[11] KOLÁŘOVÁ, J., K. FIALOVÁ, O. NOVÁČEK, M. NOVÁKOVÁ a I. PROVAZNÍK. Experimental methods for simultaneous measurement of action potentials and electrograms in isolated heart. *In: Physiological research / Academia Scientiarum Bohemoslovaca.* 1. vyd., 2012, č. 59. ISSN 0862-8408.

[12] KOZUMPLÍK, J., L. SMITAL a M. VÍTEK. *Eliminator. Software pro korekci driftu nulové izolinie signálu EKG.* Ústav biomedicínského inženýrství: Vysoké učení technické v Brně. [cit. 2012-10-28]. Dostupné z: http://www.ubmi.feec.vutbr.cz/vyzkum-a-vyvoj/produkty (software)

[13] KOZUMPLÍK, Jiří a Ivo PROVAZNÍK. *Umělá inteligence v medicíně* [online]. 2007 [cit. 2013-05-22]. Dostupné z: https://www.vutbr.cz/elearning/mod/resource/view.php?id=187774

[14] KUSUMOTO, B. F. *ECG interpretation: from pathophysiology to clinical application.* 1. vyd. New York: Springer, 2009, 107 s. ISBN 978-038-7888-798.

[15] KUTLU, Y. a D. KUNTALP. Feature extraction for ECG heartbeats using higher order statistics of WPD coefficients. *Computer metods and programs in biomedicine [online]*. 2012, č. 105, 257–267 [cit. 2012-10-27]. Dostupné z: http://www.ncbi.nlm.nih.gov/pubmed/22055998

[16] MARTIS, R. J., R. ACHARYA, A. K. RAY a Ch. CHAKRABORTY. Application of Higher Order Cumulants to ECG Signals for the Cardiac Health Diagnosis. *Proceedings of the 33rd Annual international conference of the IEEE engineering in medicine and biology society [online]*. [New York: IEEE, 2011, s. 1697-1700 [cit. 2012-11-01]. ISSN 978-1-4244-4122-8. Dostupné z: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6090487

[17] MENDEL, J. M. Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications. *Proceedings of the IEEE.* 1991, s. 278-280. DOI: 10.1109/5.75086. Dostupné z: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=75086

[18] OSOWSKI, S., L.T. HOAI a T. MARKIEWICZ. Support Vector Machine-Based Expert System for Reliable Heartbeat Recognition. *IEEE Transactions on Biomedical Engineering [online]*. 2004, roč. 51, č. 4, s. 582-589 [cit. 2012-11-07]. ISSN 0018-9294. DOI: 10.1109/TBME.2004.824138. Dostupné z: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1275573

[19] PARK, K.S., B.H. CHO, D.H. LEE, S.H. SONG, J.S. LEE, Y.J. CHEE, I.Y. KIM a S.I. KIM. Hierarchical Support Vector Machine Based Heartbeat Classification Using

Higher Order Statistics and Hermite Basis Function. *Computers in cardiology / National Institutes of Health [online]*. 2008, č. 35, 229−232 [cit. 2012-11-01]. ISSN 0276-6574. Dostupné z: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4749019&tag=1

[20] PAVLÍK, Jiří. *Aplikovaná statistika* [online]. Praha: Vysoká škola chemicko-technologická, 2005, 172 s. [cit. 2013-05-23]. ISBN 80-708-0569-2. Dostupné z: http://sofe2-files.pepiino.cz/ast/ast-vscht.pdf

[21] PETROPULU, A. *Higher-Order Spectral Analysis [online]*. CRC Press LLC, 2000 [cit. 2012-10-25]. Dostupné z: http://dsp-book.narod.ru/DSPMW/76.PD

[22] SIGG, D. C. *Cardiac electrophysiology methods and models.* New York: Springer, 2010, 491 s. ISBN 14-419-6658-7.

[23] ŠPINAR, J. *Ischemická choroba srdeční*. 1. vyd. Praha: Grada Publishing, 2003, 361 s. ISBN 80-247-0500-1.

[24] SUNG-NIEN, Y. a Ch. YING-HSIANG. Noise-tolerant electrocardiogram beat classification based on higher order statistics of subband components. *Artificial Intelligence in Medicine [online]*. 2009, č. 46, s. 165-178 [cit. 2012-28-10]. Dostupné z: http://www.ncbi.nlm.nih.gov/pubmed/19101129

[25] VÍTEK, M. a J. KOZUPMLÍK. *QRS seeker. Software pro detekci komplexů QRS.* Ústav biomedicínského inženýrství: Vysoké učení technické v Brně. (software)

[26] Wolfram MathWord: *Cumulant. [online]*. [cit. 2012-11-09]. Dostupné z: http://mathworld.wolfram.com/Cumulant.html

[27] YING-HSIANG, Ch. a Y. SUNG-NIEN. Subband Features Based on Higher Order Statistics for ECG Beat Classification. *Proceedings of the 29th Annual International Conference of the IEEE EMBS Cité Internationale [online]*. 2007, s. 1859-1862 [cit. 2012-10-28]. Dostupné z: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4352677&userType=inst

SEZNAM ZKRATEK

ANN	umělá neuronová síť (angl. artifical neutral network)
EKG	elektrokardiogram
IQR	interkvartilový interval
kap.	kapitola
SUM	normalizovaná sumace
SY	koeficient symetrie
VK	variační koeficient
VUT	Vysoké učení technické v Brně