

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

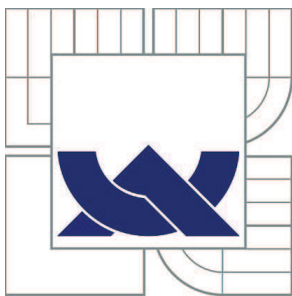
BORCENÍ ČASOVÉ OSY V OBLASTI BIOSIGNÁLŮ

DIPLOMOVÁ PRÁCE
MASTER'S THESIS

AUTOR PRÁCE
AUTHOR

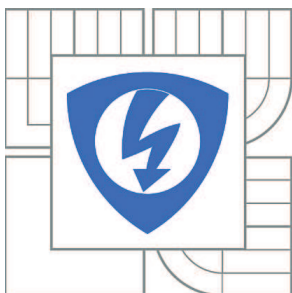
Bc. MILAN KUBÁT

BRNO 2014



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

BORCENÍ ČASOVÉ OSY V OBLASTI BIOSIGNÁLŮ

DYNAMIC TIME WARPING IN BIOSIGNAL PROCESSING

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

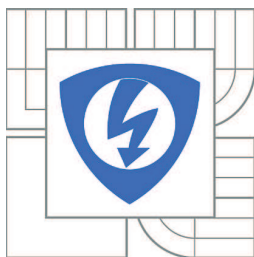
Bc. MILAN KUBÁT

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. MARTIN KLIMEK

BRNO 2014



VYSOKÉ UČENÍ
TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

Ústav biomedicínského inženýrství

Diplomová práce

magisterský navazující studijní obor
Biomedicínské a ekologické inženýrství

Student: Bc. Milan Kubát

ID: 125045

Ročník: 2

Akademický rok: 2013/2014

NÁZEV TÉMATU:

Borcení časové osy v oblasti biosignálů

POKYNY PRO VYPRACOVÁNÍ:

1) Seznamte se s metodou borcení časové osy (DTW) v oblasti biosignálů a vypracujte literární rešerši. 2) Prostudujte varianty DTW pro porovnávání signálů. 3) Navrhněte metodu pro porovnání signálů EKG metodou DTW. 4) Metodu realizujte v programovém prostředí MATLAB. 5) Realizovanou metodu borcení časové osy ověřte na reálných signálech z CSE databáze dostupné na UBMI. 6) Provedte diskuzi získaných výsledků.

DOPORUČENÁ LITERATURA:

[1] KEOGH, E.J., PAZZANI, M.J.: Derivative Dynamic Time Warping. First SIAM International Conference on Data Mining, Chicago, 2001, pp. 285–289.

[2] ACKENHUSEN, J. G.: Real-time signal processing. Prentice Hall, New Jersey, 1999. ISBN 978-0136317715.

Termín zadání: 10.2.2014

Termín odevzdání: 23.5.2014

Vedoucí práce: Ing. Martin Klimek

Konzultanti diplomové práce:

prof. Ing. Ivo Provazník, Ph.D.

Předseda oborové rady

UPOZORNĚNÍ:

Autor diplomové práce nesmí při vytváření diplomové 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.

ANOTACE

Práce je věnována metodě borcení časové osy v oblasti biosignálů, zejména pak jejímu využití pro signály EKG. V úvodu jsou shrnuty teoretické poznatky o elektrokardiografii, následuje rozbor metody DTW společně se stanovením podmínek a požadavků na úspěšnou aplikaci algoritmu. Jsou popsány varianty metody a možnosti její aplikace. Náplní části praktické je návrh samotné metody, seznámení se s výstupy, optimalizace nastavení a realizace metod s borcením časové osy souvisejících.

KLÍČOVÁ SLOVA

EKG, DTW, DDTW, borcení časové osy, biosignál, kumulovaná vzdálenost, euklidovská vzdálenost, matice vzdáleností, po částech lineární aproximace, shluková analýza, segmentace

ABSTRACT

This work is dedicated to dynamic time warping in biosignal processing, especially it's application for ECG signals. On the beginning the theoretical notes about cardiography are summarized. Then, the DTW analysis follows along with conditions and demands assessments for it's successful application. Next, several variants and application possibilities are described. The practical part covers the design of this method, the outputs comprehension, settings optimization and realization of methods related with DTW

KEYWORDS

ECG, DTW, DDTW, Dynamic Time Warping, biosignal, cumulated distance, Euclidean distance, distance matrix, piecewise linear approximation, cluster analysis, segmentation

KUBÁT, M. *Borcení časové osy v oblasti biosignálů*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2014. 59 s. Vedoucí diplomové práce Ing. Martin Klimek.

Prohlášení

Prohlašuji, že jsem tuto diplomovou práci na téma *Borcení časové osy v oblasti biosignálů* vypracoval samostatně pod vedením vedoucího semestrálního projektu 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 uvedené diplomové práce prohlašuji, že v souvislosti s vytvořením této práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení § 152 trestního zákona č. 140/1961 Sb.

V Brně dne 23. května 2014

.....

podpis autora

Poděkování

Děkuji Ing. Martinu Klimkovi za vypsání zajímavého tématu, za možnost vypracování diplomové práce pod jeho vedením a rady poskytnuté k jejímu vypracování.

V Brně dne 23. května 2014

.....

podpis autora

Obsah

ÚVOD	7
1 ÚVOD K BIOLOGICKÝM SIGNÁLŮM A EKG.....	8
1.1 EKG	8
1.2 PŘEDZPRACOVÁNÍ EKG.....	9
1.2.1 Odstranění rušení	10
1.2.2 Rozměření signálu	10
2 BORCENÍ ČASOVÉ OSY	12
2.1 ALGORITMUS METODY	12
2.1.1 Vlastnosti optimální funkce.....	13
2.1.2 Omezení vyhledávacího algoritmu.....	15
2.1.3 Výpočet vzdálenosti.....	18
2.1.4 Váhová funkce, normalizační faktor.....	19
2.2 DDTW.....	20
3 VYUŽITÍ DTW PRO SIGNÁLY EKG	22
3.1 DTW JAKO PŘEDSTUPEŇ SHLUKOVÉ ANALÝZY	22
3.2 AUTOMATIZOVANÁ SEGMENTACE EKG	24
3.3 PŘÍKLADY JINÝCH VYUŽITÍ.....	26
4 PROGRAMOVÉ REALIZACE ALGORITMŮ.....	27
4.1 POUŽITÉ SIGNÁLY.....	27
4.1.1 Filtrace.....	27
4.1.2 Rozměření	28
4.2 DTW	29
4.2.1 Vliv rozdílů v porovnávaných signálech.....	32
4.2.2 Omezení	39
4.2.3 DDTW	41
4.2.4 Shrnutí.....	42
4.3 SHLUKOVÁ ANALÝZA.....	43
4.4 EKG SEGMENTACE.....	44
4.4.1 PLA	45
4.4.2 Vytvoření vzorů	45
4.4.3 Testování vytvořeného algoritmu	47
ZÁVĚR	48
POUŽITÉ ZDROJE	49
SEZNAM OBRÁZKŮ.....	51
SEZNAM ZKRATEK	53
OBSAH PŘILOŽENÉHO CD	54
DODATKY	55

Úvod

Cílem diplomové práce je prostudovat metodu borcení časové osy v oblasti biologických signálů. Pochopit její princip, poznat varianty algoritmu, navrhnout metody pro porovnání signálů EKG a realizovat je v programovém Prostředí MATLAB.

Lidské srdce během života podává neúnavný, nepřetržitý výkon. Protože jeho nesprávná funkce může významně ovlivnit kvalitu života, byly vyvinuty rozličné metody pro jeho diagnostiku. Elektrokardiogram popisuje jak srdeční rytmus, tak způsob šíření vzruchu v srdci, proto se elektrokardiografie stala nedílnou součástí klinické praxe a zároveň jednou z nejcennějších diagnostických metod.

Na základě vyhodnocení průběhu zaznamenaného napětí srdečního cyklu lze usuzovat o jejich normalitě či abnormalitě a ruku v ruce s tím o stavu srdečního svalu. Současná „digitalizovaná“ doba přináší takřka nepřeborné možnosti analýzy EKG záznamů při využití mnoha různých algoritmů.

Člověk má stále v diagnostice nemocí nezastupitelné postavení a ještě dlouho jej pravděpodobně mít bude. V mnoha ohledech je však technika schopná alespoň podat pomocnou ruku.

Například zdlouhavé prohlížení záznamů může jednak být unavující, ale hlavně odvádí diagnostika od jiných povinností. Zde se kupříkladu nabízí možnost klasifikovat jednotlivé cykly do reprezentativních skupin, což časovou náročnost sníží. Problémem takové klasifikace je proměnlivost jak srdečního rytmu, tak průběhu a úrovně vln a kmitů, kdy se vyhodnocení vzájemné podobnosti komplikuje. K úspěšnému porovnání je nezbytný nelineární algoritmus. Metoda borcení časové osy, která je s úspěchem využívána například i u rozpoznávání zvukových signálů, takovým algoritmem je.

Jedná se o metodu využitelnou v řadě aplikací porovnávajících a klasifikujících průběhy EKG záznamu na základě nelineární úpravy časové osy signálu za účelem dosažení odpovídajícího zarovnání a přesnějšího výpočtu (ne)podobnosti (vzdálenosti).

Práce bude rozdělena na teoretickou a praktickou část. V rámci části teoretické bude po stručném zopakování základů elektrokardiografie a zpracování signálů představena metoda borcení časové osy. Následující kapitola obě předcházející spojí a prezentuje možnosti borcení časové osy pro signály EKG. Část praktická bude věnována programové realizaci navržených metod a ověření závěrů části teoretické.

1 Úvod k biologickým signálům a EKG

Pojem biologický signál (biosignál) zahrnuje projevy fyzikálních polí vznikajících v důsledku činností jednotlivých orgánů. Tyto signály mohou být

- elektrické (EKG, EEG, EOG, EGG, ...),
- magnetické (MKG, MEG, MRG, ...),
- mechanické (pletysmogram, pneumotachogram, ...),
- akustické (FKG, ...), apod.

Mají diagnostickou hodnotu, jejich zaznamenáním a vyhodnocením hodnot lze získat informace o stavu celého organismu nebo příslušných orgánů [7].

1.1 EKG

Srdce je ústředním orgánem krevního oběhu. Funguje jako pumpa vytlačující krev do cév celého těla. Je tvořeno zvláštním typem svaloviny, srdeční svalovinou, jejíž buňky jsou navzájem spojeny. Anatomicky je rozděleno na levou a pravou srdeční síň a levou a pravou srdeční komoru.

Srdce obsahuje dva typy vláken:

- **převodní systém** – obsahuje buňky schopné spontánní depolarizace, zde vzniká podráždění, a
- **pracovní myokard** – reaguje stahy na vzruchy vytvořené převodním systémem.

Vzruch vzniká v sinoatriálním uzlu (SA uzel), kde dochází k samovolným elektrickým podrážděním. Vznik vzruchů v sinoatriálním uzlu je spojen s tepovou frekvencí pohybující se v rozmezí 60 až 80 tepů za minutu. Podráždění se šíří do atrioventrikulárního uzlu (AV uzel), za kterým následuje Hissův svazek. Ten se rozbíhá v levé a pravé Tawarovo raménko. Raménka se v komorách dělí na Purkyňova vlákna.

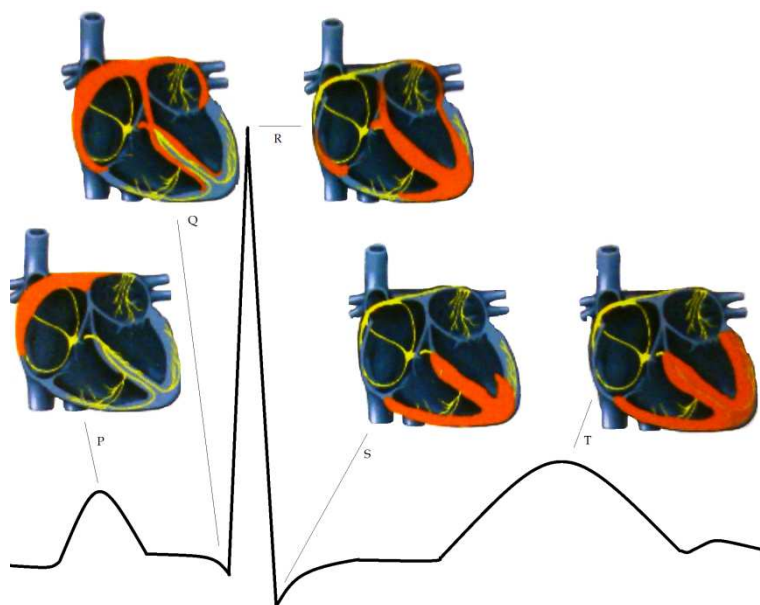
EKG, elektrokardiogram, je záznamem elektrické aktivity srdečního svalu. Jedná se o takzvané akční potenciály srdeční. Vzruchy přenášené převodním systémem jsou tak malé, že jsou prakticky nezaznamenatelné. Měřené rozdíly potenciálů jsou způsobeny až reakcí pracovního myokardu.

V záznamu EKG se objevují významné kmity související s průběhem akčního napětí v srdečním svalu. Velikost, tvar a poloha těchto kmitů jsou důležitým diagnostickým prostředkem pro lékaře. Vhodným rozměřením lze určit tepovou frekvenci, některé změny vyvolané nedostatečným prokrvením myokardu, srdeční infarkt, i další poruchy [11].

- **Vlna P** odpovídá stahu (depolarizaci) síní.
- **Komplex QRS**, nejvýraznější úsek signálu, odpovídá depolarizaci komor. QRS komplex zároveň překrývá vlnu, která by odpovídala repolarizaci síní.

- **Vlna T** je známkou repolarizace komor.

Souvislost srdeční aktivity se signálem EKG je znázorněna na obrázku (Obrázek 1). Šíření vzruchu v srdci je zde naznačeno červenou barvou.

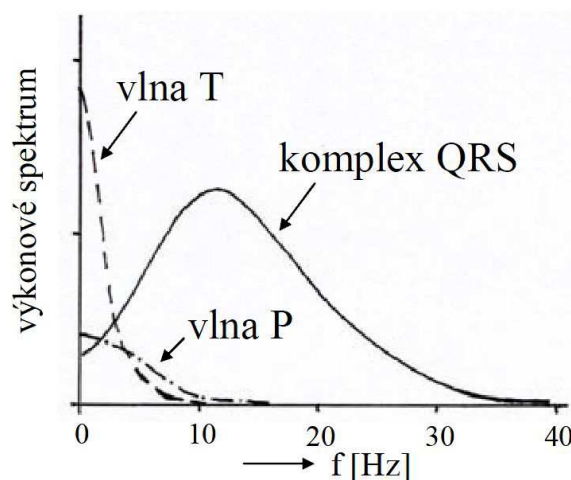


Obrázek 1 Souvislost srdeční aktivity se signálem EKG. Podle [11], upraveno.

1.2 Předzpracování EKG

Blok předzpracování je při práci se signálem zařazen za účelem úpravy signálu do podoby vhodné z hlediska následné analýzy. Nežádoucí složky obsažené v sejmutém signálu bývají nejčastěji odstraňovány filtrací. Při aplikaci filtrů musíme dbát na vlastnosti užitečného signálu (například ty spektrální), aby nedošlo k jeho zkreslení [7].

Frekvenční pásmo klidového EKG se pohybuje do 125 Hz. Hrubý náčrt výkonových spekter vlny T, P a komplexu QRS je na obrázku níže (Obrázek 2).



Obrázek 2 Náčrt spekter základních vln a kmitů EKG

1.2.1 Odstranění rušení

Snímání signálu EKG provází několik šumových složek typických i pro biosignály jiné. Zaprvé je nutno brát v úvahu fakt, že srdce není jediným zdrojem biosignálu v lidském těle. I samotný způsob snímání, popřípadě přístroj ke snímání používaný mohou do výsledku zavést rušivé vlivy. Při vyhodnocování záznamů se pracovníci potýkají nejčastěji s rušením myopotenciály, kolísáním nulové izolinie a síťovým brumem [7].

- Rušení **myopotenciály** je způsobeno svalovou činností. Jedná se o širokopásmový šum. Při snímání v klidu zabírají pásmo od 100 Hz výše. Ovšem u signálů snímaných při zátěži, jako například při ergometrii, začínají už na 10 až 20 Hz.
- **Kolísání nulové izolinie**, takzvaný drift, jehož frekvenční spektrum sahá maximálně po 2 Hz, je způsobený elektrochemickými procesy na rozhraní elektroda - kůže, dýcháním a pomalými pohyby pacienta. Běžně se jeho vliv omezuje používáním vhodných elektrod nebo například speciálních EKG gelů snižujících přechodové odpory.
- **Síťovým brumem** je nazýváno rušení způsobené indukovanou nebo kapacitní vazbou od střídavého proudu z elektrorozvodné sítě. Síťové rušení je téměř harmonické, úzkopásmové, pohybující se kolem 50 Hz.

Vysokofrekvenční rušení potlačíme filtrem typu dolní propust s mezní frekvencí 70 popřípadě 80 Hz [3][10]. V [3] je jako příklad uvedena dolní propust druhého řádu Lynnova typu. Síťové rušení odstraňují úzkopásmové zádrže se střední frekvencí na 50 Hz. Drift lze potlačit filtrem založeným na principu nulování spektrálních čar [10].

Bližší rozbor filtrace přesahuje rámec této diplomové práce, odkazuji zde proto na svou bakalářskou práci, viz [12].

1.2.2 Rozměření signálu

Po odstranění případných šumových složek je pro tuto práci dále nutné ze signálu extrahovat jednotlivé cykly. Prvním krokem je obvykle detekce QRS komplexu. Detektorů, které označí pozici R-vlny, existuje velké množství. Příkladem mohou být detektory založené na umocnění filtrovaného signálu, na obálce filtrovaného signálu popřípadě na rozkladu signálu bankou filtrů.

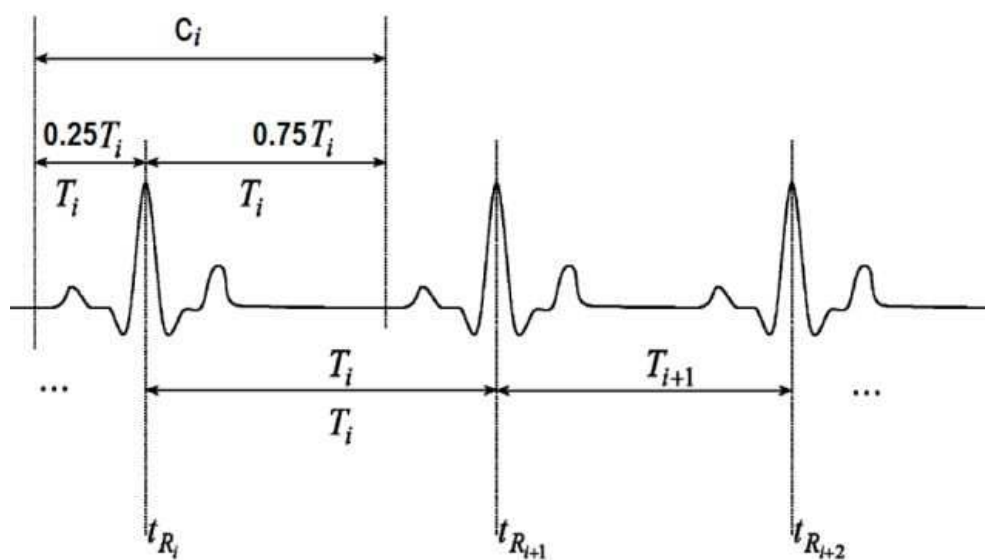
Spolehlivější detekci než jednodušší detektory uvedené v předchozím odstavci poskytuje detektor QRS založený na vlnkové transformaci [9].

Zjednodušeně lze z pozic jednotlivých R vln

$$t_R = \{t_{R1}, t_{R2}, \dots, t_{Rn}\}, \quad (1)$$

odvodit jejich vzdálenost T_i . Extrakci samotných cyklů nyní provedeme tak, že 75 % vzdálenosti R-vln přiřadíme aktuálnímu cyklu, zbylých 25 % cyklu následujícímu. Blíže tento postup rozvádí následující obrázek (Obrázek 3), kde C_i je konkrétní i-tý cyklus EKG [10].

Tato metoda však nemusí být příliš spolehlivá a možnost jejího využití je nezbytné předem prověřit na vybraném souboru signálů.



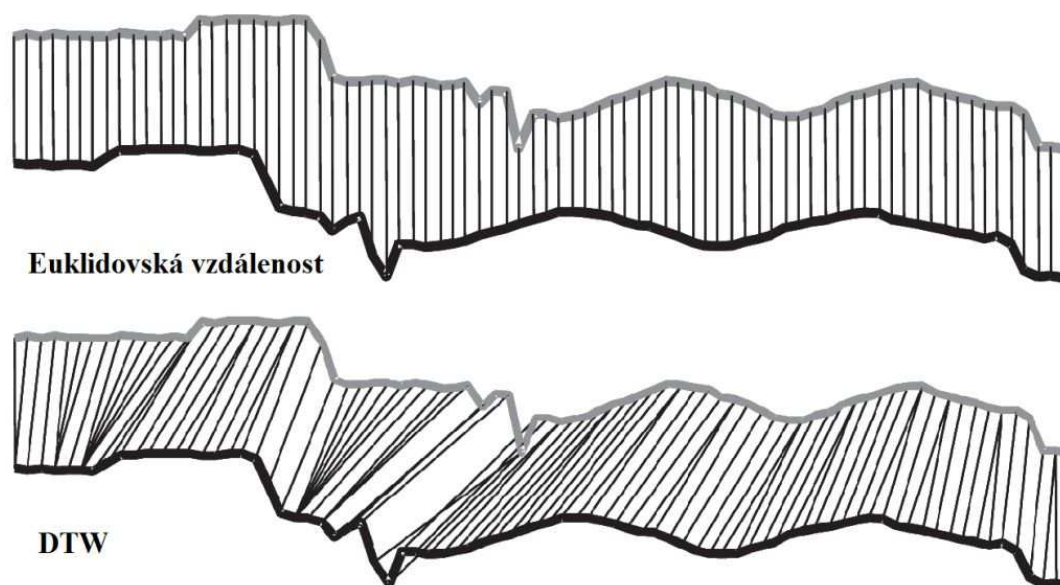
Obrázek 3 Příklad rozměření signálu EKG [10].

2 Borcení časové osy

V praxi se běžně setkáváme s prezentací dat v podobě časové posloupnosti hodnot. V oblasti biomedicínského inženýrství se převážně jedná o záznamy biologických signálů, až už srdce, svalů či mozku. Častým úkolem bývá porovnání jedné sekvence s jinou. V široké škále oborů lze podobnost hodnotit na základě prostého měření Euklidovských vzdáleností. Mnohdy se však setkáváme se signály s podobným průběhem, avšak jiným rozmístěním odpovídajících si hodnot na časové ose. EKG je toho zářivým příkladem [1].

Odpověď při hledání podobností mezi signály, při jejich zarovnávání či předzpracování přináší algoritmus na základě dynamického programování, kdy musíme „zbortit“ časovou osu jednoho nebo obou signálů. Nazývá se Dynamic Time Warping (DTW, „borcení časové osy“) Tento algoritmus pracuje s efektem nelineární časové normalizace, kde kolísání v časové ose je modelováno „bortivou“ (DTW) funkcí se specifikovanými vlastnostmi [1].

Rozdíl mezi určováním Euklidovské vzdálenosti a vzdálenosti u metody DTW ilustruje následující obrázek (Obrázek 4).



Obrázek 4 Využití Euklidovské vzdálenosti a DTW. Převzato z[6], upraveno.

2.1 Algoritmus metody

Pro ilustraci uvažujme dvě časové závislosti hodnot, signály, délky m respektive n

$$\begin{aligned} A &= \{a_1, a_2, \dots, a_i, \dots, a_m\} \quad \text{a} \\ B &= \{b_1, b_2, \dots, b_j, \dots, b_n\}, \end{aligned} \tag{2}$$

kde a_i je i -tý prvek signálu A a b_j je j -tý prvek signálu B.

Sestrojíme $m \times n$ matici lokálních vzdáleností mezi hodnotami a_i a b_j . Každý prvek matice (i,j) tedy odpovídá zarovnání těchto hodnot. Cílem je nalézt podle stanovených pravidel optimální cestu

$$j = \psi(i) \quad (3)$$

minimalizující funkci D celkové vzdálenosti mezi signály A a B

$$DTW(A, B) = \sum_{n=1}^I \delta[a(j), b(\psi(i))]. \quad (4)$$

Jedná se o součet lokálních vzdáleností mezi i -tým vzorkem signálu A a j -tým signálu B.

Zavedením obecné časové proměnné k a vyjádřením časových proměnných i a j jako funkcí k , dostáváme jednoduché závislosti.

$$\begin{aligned} i &= r(k), & k &= 1, \dots, K, \\ j &= s(k), & k &= 1, \dots, K, \end{aligned} \quad (5)$$

zde K je celková délka obecné časové osy při porovnávání signálů A a B [2].

Výslednou cestu W lze také chápat jako posloupnost prvků matice vzdáleností mezi A a B [1]. Označíme-li k -tý prvek cesty jako $w_k = (i, j)_k$, dostáváme

$$\begin{aligned} W &= \{w_1, w_2, \dots, w_k, \dots, w_K\} \\ \max(m, n) &\leq K < m + n - 1. \end{aligned} \quad (6)$$

2.1.1 Vlastnosti optimální funkce

Předpokladem smysluplnosti výsledné nalezené cesty, optimální funkce, je splnění několika požadavků.

Okrajové podmínky

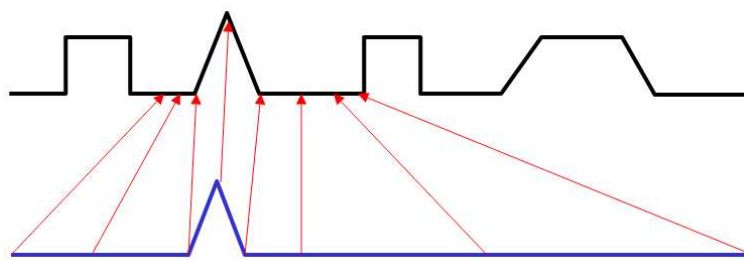
Podmínky

$$w_1 = (1,1) \text{ a } w_K = (m,n); [1], \quad (7)$$

případně

$$\begin{aligned} r(1) &= 1, & s(1) &= 1, \\ r(K) &= m, & s(K) &= n; [2] \end{aligned} \quad (8)$$

vyjadřují požadavek, aby výsledná cesta začínala a končila v protilehlých rozích matice vzdáleností. Pokud by to tak nebylo, algoritmus by mohl ignorovat část jedné sekvence a zarovnání by bylo bezvýznamné [14].



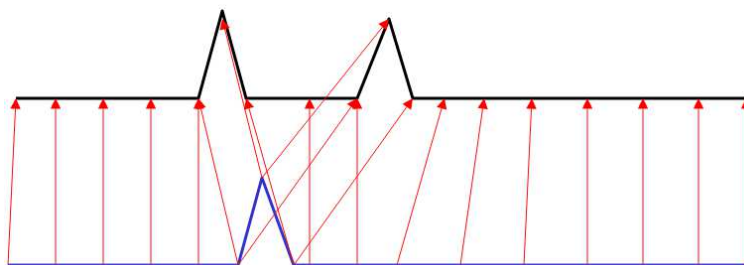
Obrázek 5 Příklad nesplnění okrajových podmínek [14].

Monotónnost

Výsledná cesta musí být neklesající, aby nebyly brány v potaz hodnoty již jednou započítané. Tedy, aby byla zachována časová posloupnost hodnot a algoritmus nevkládal hodnoty „z minulosti“. Předchází-li hodnotu $w_k = (a, b)$ hodnota $w_{k-1} = (a', b')$, musí platit

$$a - a' \geq 0 \quad a \quad b - b' \geq 0; [6]. \quad (9)$$

Pokud by se v jedné posloupnosti opakoval určitý rys dvakrát, algoritmus by se pokusil toto vysvětlit hodnotami z minulosti a opět by vyprodukoval nesmyslné řešení.



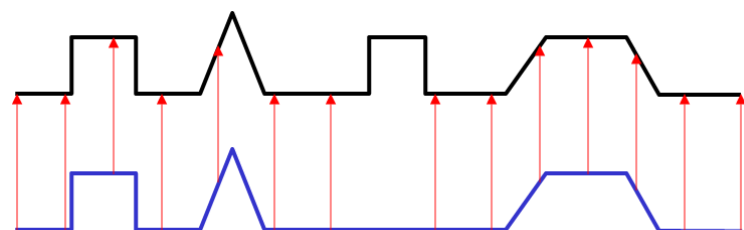
Obrázek 6 Následek nemonotónní výsledné cesty [14].

Spojitosť

Budeme-li uvažovat dvě po sobě jdoucí hodnoty z předchozího příkladu, musí také platit

$$a - a' \leq 1 \quad a \quad b - b' \leq 1; [6]. \quad (10)$$

Toto omezení, které by se dalo spojit do jednoho vztahu s omezením předešlým, předchází tomu, aby algoritmus DTW při porovnávání vynechal některé segmenty signálu [2][14].

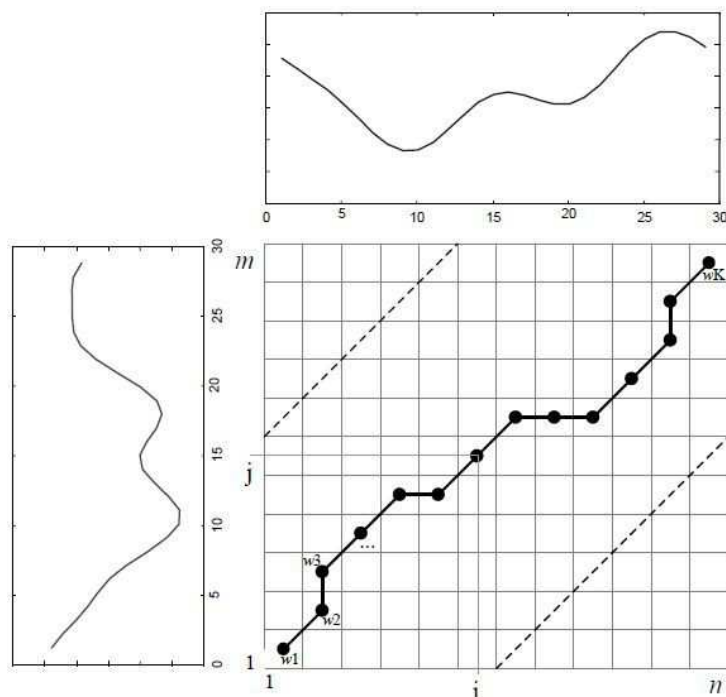


Obrázek 7 Následek nespojitosti výsledné cesty [14].

Výslednou cestu poté lze účinně nalézt pomocí dynamického programování, kdy sledujeme kumulativní vzdálenost $\gamma(i, j)$, definovanou rekurzivně jako vzdálenost $\delta(i, j)$ v aktuálním prvku matice a minimální kumulativní vzdálenost v přilehlých prvcích [1]. Čili

$$\gamma(i, j) = \delta(i, j) + \min \begin{cases} \gamma(i-1, j-1) \\ \gamma(i, j-1) \\ \gamma(i-1, j) \end{cases} . \quad (11)$$

Průběh bortivé cesty může vypadat například takto (Obrázek 8).



Obrázek 8 Příklad bortivé cesty [1].

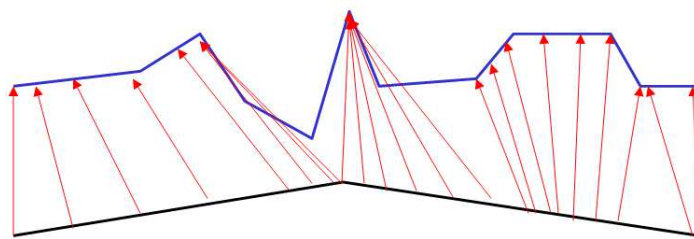
2.1.2 Omezení vyhledávacího algoritmu

Hledání výsledné cesty zpravidla podléhá několika omezením, které ho směřují do smysluplné oblasti matice vzdáleností. Těmito omezeními lze zároveň částečně snížit výpočetní náročnost. Na druhou stranu je v mnoha aplikacích nesnadné nalézt ideální podobu těchto omezení [1][2][6].

Omezení strmosti

Je zřejmé, že výsledná funkce DTW by neměla obsahovat příliš velký ani příliš malý přírůstek. Nadměrná komprese nebo naopak dilatace časové osy může být důvodem nesprávné korespondence například mezi krátkým úsekem signálu A a relativně dlouhým úsekem signálu B.

Omezení strmosti funkce můžeme realizovat konkrétně tak, že pokud se při rostoucím k pohybuje funkce ve směru jedné osy po sobě n -krát, není jí po určitý počet kroků nadále dovoleno se tímto směrem pohybovat. Tento směr je opět možný až po „odkrokování“ m kroků v jiném směru [2].



Obrázek 9 Následek mnoha kroků ve vertikálním nebo horizontálním směru [14].

Další možností je nahrazení rovnice (11) rovnicí

$$\gamma(i, j) = \delta(i, j) + \min \begin{cases} \gamma(i-1, j-1) \\ X\gamma(i-1, j) \\ X\gamma(i, j-1) \end{cases}, \quad (12)$$

kde X je kladné reálné číslo. Zvětšováním hodnoty X je hledání výsledné cesty směřováno diagonálně [6].

Směr kroku

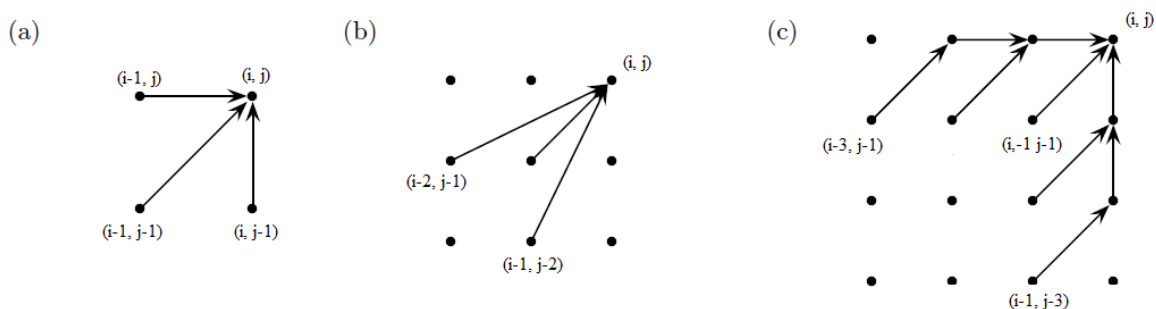
Rovnice popisující výpočet mohou být znázorněny jako schémata kroků přípustných v každé fázi hledání. Kromě rovnice (10) to může být například i rovnice

$$\gamma(i, j) = \delta(i, j) + \min \begin{cases} \gamma(i-1, j-1) \\ \gamma(i-1, j-2), \\ \gamma(i-2, j-1) \end{cases} \quad (13)$$

takto se při každém kroku hledání musíme také jednou posunout v diagonálním směru [1]. Nebo vztah (podle [19], upraveno do formátu předchozích rovnic)

$$\gamma(i, j) = \delta(i, j) + \min \begin{cases} \gamma(i-1, j-1) \\ \gamma(i-2, j-1) + \delta(i-1, j) \\ \gamma(i-1, j-2) + \delta(i, j-1) \\ \gamma(i-3, j-1) + \delta(i-2, j) + \delta(i-1, j) \\ \gamma(i-1, j-3) + \delta(i, j-2) + \delta(i, j-1) \end{cases} \quad (14)$$

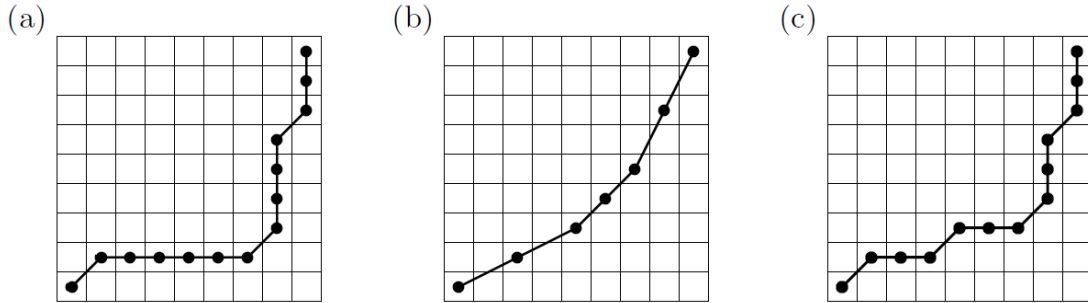
Grafická znázornění kroků podle uvedených rovnic jsou v následujících schématech (Obrázek 10) [19].



Obrázek 10 Možné směry pohybu funkce (a) pro rovnici 11, (b) pro 13 a (c) pro 15 [19].

Obrázek 11 (a) znázorňuje, k jakým degeneracím výsledné cesty může při využití rovnice 11 docházet. Při využití rovnice 13 může při zarovnání docházet k vynechání elementů hodnocených signálů (b). (c) předchází vynecháním a přidává váhu šikmému pohybu.

V dalším textu budou kroky pro zjednodušení označovány jako (a), (b) a (c).



Obrázek 11 Tři bortivé cesty s ohledem na rozdílný krok [19].

Vzhledem k okolí, ze kterého je vypočítávána kumulovaná vzdálenost, je před vlastním výpočtem nutno matici rozšířit o řádky a sloupce s velmi vysokými hodnotami, ideálně nekonečnem. V případě (b) o jeden, v případě (c) o dva řádky a sloupce [20]. Samozřejmě existuje velké množství dalších různých směrů pohybu, viz například v [2].

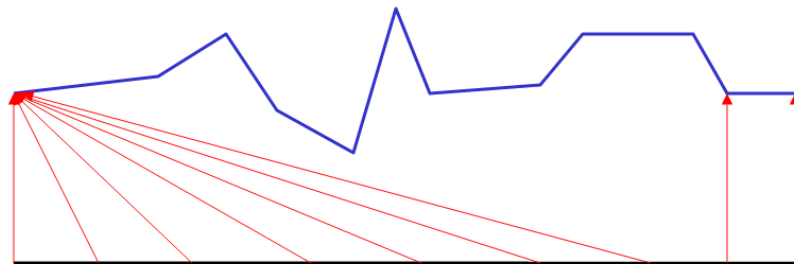
Globální omezení pohybu funkce, metoda okna

Zobecněním lokálních omezení na celou matici lze vymezit oblast přípustnou pro pohyb funkce DTW jako

$$\begin{aligned} 1 + \alpha[r(k) - 1] &\leq s(k) \leq 1 + \beta[r(k) - 1], \\ n + \beta[r(k) - m] &\leq s(k) \leq n + \alpha[r(k) - m], \end{aligned} \quad (15)$$

kde α a β jsou minimální respektive maximální směrnice přímky vymezující přípustnou oblast [2].

Další omezení oblasti pohybu funkce DTW vychází z předpokladu, že uvažované signály nemohou obsahovat zásadní časové rozdíly mezi jednotlivými úseky [6]. Je-li výsledná cesta příliš vzdálená od diagonály matice, algoritmus přeskakuje různé rysy a zasekává se na těch podobných [14].

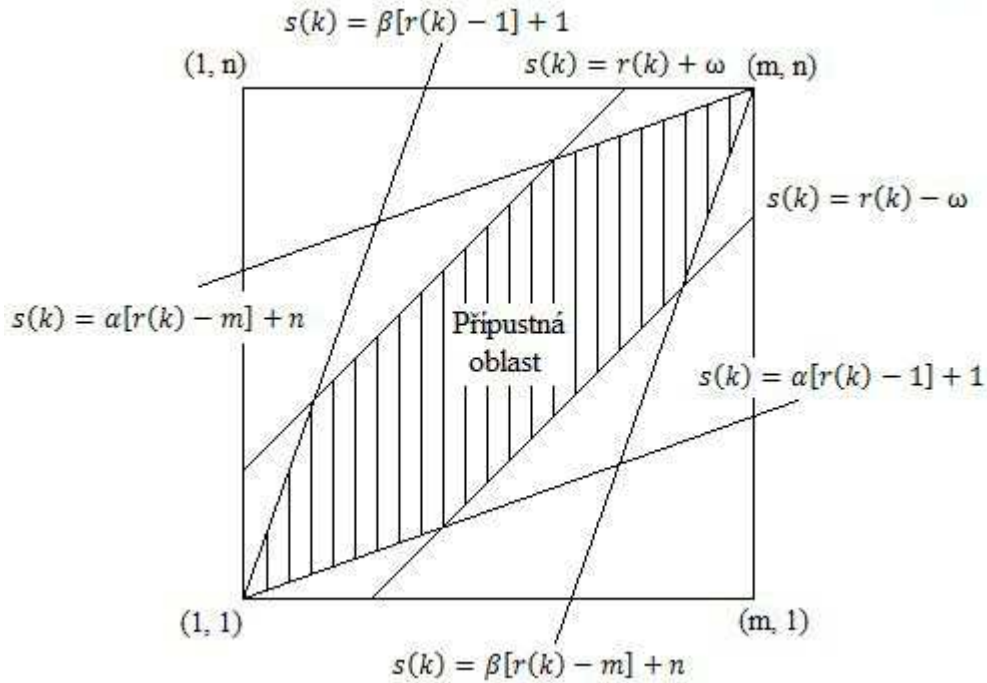


Obrázek 12 Následek pohybu funkce daleko od diagonály[14].

Z tohoto důvodu lze stanovit podmínku

$$|r(k) - s(k)| \leq \omega, \quad (16)$$

kde ω je vhodné kladné celé číslo někdy nazývané jako šířka okna.



Obrázek 13 Příklad omezujících podmínek. Podle [2].

Omezení (15) má tu nevýhodu, že šířka okna musí být větší než rozdíl $|m - n|$, aby bylo do globálně omezené oblasti možno zahrnout koncový bod. Tento nedostatek lze odstranit použitím vztahu

$$\left| \frac{n}{m} r(k) - s(k) \right| \leq \omega. \quad (17)$$

Případ globálních vymezení je znázorněn v předchozím obrázku (Obrázek 13), kde je i označena přípustná oblast pohybu funkce DTW [2].

2.1.3 Výpočet vzdálenosti

Jako měřítko vzdálenosti může dostatečně posloužit prostá Euklidovská vzdálenost mezi dvěma body v jednorozměrném prostoru

$$\delta(a_i, b_j) = |a_i - b_j|. \quad (18)$$

Samozřejmostí je možnost využití jakékoli jiné vzdálenosti [2].

Obecný tvar pro určení skutečné celkové minimální vzdálenosti mezi signály A a B lze zapsat jako

$$DTW(A, B) = \min \left[\frac{\sqrt{\sum_{k=1}^K w_k}}{K} \right], \quad (19)$$

kde je K ve jmenovateli použito ke kompenzaci faktu, že výsledné cesty mohou mít různé délky [6].

Dalším možným vyjádřením je

$$DTW(A, B) = \min_{\{r(k), s(k), K\}} \left[\frac{\sum_{k=1}^K \delta[r(k), s(k)] \hat{W}(k)}{N(\hat{W})} \right], \quad (20)$$

kde $\hat{W}(k)$ je hodnota váhové funkce pro k -tý úsek funkce DTW a $N(\hat{W})$ je normalizační faktor, který je funkcí váhové funkce [2].

2.1.4 Váhová funkce, normalizační faktor

Tato podkapitola přibližuje význam veličin uvedených ve vztahu (18). Váhová funkce ve své podstatě koresponduje s odstavcem „Směr pohybu funkce“ kapitoly 2.1.2

Váhová funkce $\hat{W}$ závisí jen na lokální cestě. Existuje více typů, z nichž nejčastější jsou následující [2].

Typ a) symetrická váhová funkce

$$\hat{W}(k) = [r(k) - r(k - 1)] + [s(k) - s(k - 1)], \quad (21)$$

typ b) asymetrická váhová funkce

$$\text{b1)} \quad \hat{W}(k) = r(k) - r(k - 1), \quad (22)$$

$$\text{b2)} \quad \hat{W}(k) = s(k) - s(k - 1), \quad (23)$$

typ c)

$$\hat{W}(k) = \min[r(k) - r(k - 1), s(k) - s(k - 1)], \quad (24)$$

typ d)

$$\hat{W}(k) = \max[r(k) - r(k - 1), s(k) - s(k - 1)]. \quad (25)$$

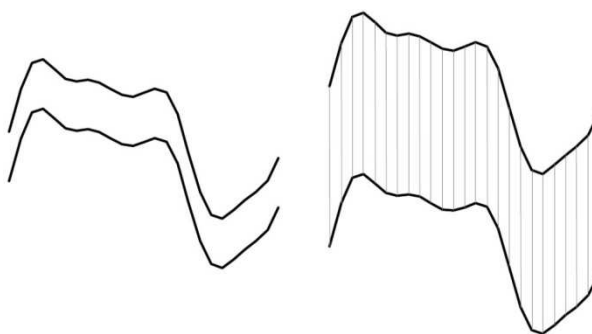
Podobně jako je v rovnici (17) použita hodnota K ke kompenzaci různých délek možných cest, zde je definován takzvaný normalizační faktor

$$N(\hat{W}) = \sum_{k=1}^K \hat{W}(k). \quad (26)$$

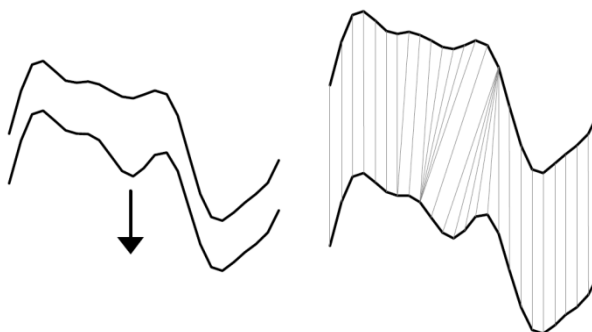
2.2 DDTW

Přestože je algoritmu DTW s úspěchem využito v mnoha aplikacích, obdržené výsledky mohou být nežádoucí. Klíčovým faktem je, že algoritmus se může pokusit vysvětlit proměnlivost v y-ové ose borcením osy x-ové. Toto může vést k neintuitivním zarovnáním, kdy je jeden bod signálu ztotožněn s mnoha body signálu druhého. Tyto případy se nazývají jako singularity.

DTW dále nemusí nalézt zřejmé zarovnání dvou posloupností jednoduše proto, že určitý rys (vrchol, inflexní bod, atd.) jedné posloupnosti je mírně vyšší či nižší než odpovídající rys v posloupnosti druhé. V prvním příkladu (Obrázek 14) jsou dvě identické posloupnosti a jejich zarovnání. Změníme-li v jednom úseku hodnoty posloupnosti, algoritmus na tuto změnu reaguje borcením osy časové, čímž vytvoří dvě singularity (Obrázek 15).



Obrázek 14 Zarovnání dvou identických posloupností [1]



Obrázek 15 Vznik singularit při mírné změně [1]

Za účelem předcházení singularitám bylo zavedeno mnoho přístupů jako například omezení uvedená v kapitole 2.1.2. Tyto metody však mohou znemožnit nalezení správné bortící funkce. Dále není úplně zřejmé, jak nastavit příslušné parametry těchto omezení [1].

Další možností je variace metody DTW nazvaná Derivative Dynamic Time Warping (DDTW). Využitím první difference lze rozlišit případy, kdy jsou dva body stejné hodnoty součástí opačných trendů, jeden trendu klesajícího, druhý rostoucího.

Algoritmus metody je v podstatě shodný s algoritmem DTW. Jednotlivé prvky vzdálenostní matice však určíme jako čtverec rozdílu odhadu diferencí

$$\delta(a', b') = (a' - b')^2. \quad (27)$$

Odhad diferencí můžeme obecně určit jako rozdíl dvou sousedních hodnot. Vzhledem k odlehlým hodnotám je ale robustnější následující vztah

$$a'_i = \frac{(a_i - a_{i-1}) + \frac{(a_{i+1} - a_{i-1})}{2}}{2} . \quad (28)$$

Ve své podstatě se jedná o průměr směrnice přímký spojující daný bod a bod předcházející a přímký spojující sousední body daného bodu. Za povšimnutí stojí, že tento odhad není definován pro první a poslední bod posloupnosti [1].

3 Využití DTW pro signály EKG

Vzhledem k charakteru metody borcení časové osy se ve většině aplikací jedná o vzájemné porovnávání jednotlivých cyklů záznamu, popřípadě o porovnávání cyklů s určitou referencí. Na základě výsledků pak obecně dochází ke klasifikaci. Jelikož porovnáním cyklů EKG přímo obdržíme míru jejich podobnosti (vzdálenost), jako další krok v analýze signálu se přímo nabízí shluková analýza. Také je možné vyhledávat v signálech typické průběhy (ať už normální či abnormální průběhy).

Metodou DTW získáme matici vzdáleností jednotlivých cyklů EKG, aplikací některé metody shlukování lze tyto cykly rozdělit na normální a abnormální. Zároveň lze vybrat reprezentativní cyklus daného shluku. Vytříděním reprezentativních cyklů zkrátíme časovou náročnost často vyžadované lékařské kontroly záznamu [10].

3.1 DTW jako předstupeň shlukové analýzy

Diagnostika EKG spadá a pravděpodobně ještě po dlouhou dobu spadat bude do kompetence lékařů. Výpočetní technika však může být užitečným nástrojem pro usnadnění práce zdravotníka a pro úsporu jeho času. Při dlouhodobém elektrokardiografickém monitoringu trvajícím zpravidla 24 až 48 hodin by bylo při běžné tepové frekvenci nutné projít řádově statisíce cyklů. Z hlediska časové náročnosti je takový přístup nerealizovatelný. Automatizovaná klasifikace cyklů do tříd na základě stanoveného stupně vzájemné (ne)podobnosti umožňuje lékařovi vybrat reprezentativní vzorek z každé třídy a usuzovat o dané diagnóze.

Shluková analýza je souhrnné označení skupiny matematických metod používaných ke zjištění, které objekty z dané sady jsou podobné. Podobné objekty jsou zařazeny do stejného shluku. Tato metoda má široké využití ve velké škále oborů, jelikož na klasifikaci je založeno mnoho výzkumů, ať už medicínských, tak i hydrologických či ekonomických [12].

Metody shlukové analýzy lze rozdělit na nehierarchické a hierarchické. První jmenované jsou typické rozkladem množiny objektů na předem daný počet shluků bez vytvoření hierarchické struktury. Tento přístup je pro potřeby práce nevhodný. Výstupem hierarchických metod jsou hierarchicky uspořádané shluky. Rozlišujeme shlukování aglomerativní, kde jsou jednotlivé objekty postupně shlukovány, a shlukování divizní, kde algoritmus postupuje od jediného shluku, který je rozkládán na shluky menší. Vzhledem ke vhodnosti použití budou dále využity hierarchické aglomerativní metody.

Algoritmus zahrnuje několik základních kroků, přičemž algoritmus DTW pokrývá první tři kroky metody (získání matice dat, její standardizace a samotný výpočet vzdáleností objektů), jelikož přímo poskytuje vzdálenosti objektů (cyklů).

Jak bylo zmíněno, vycházíme z jednotlivých cyklů a známých vzdáleností mezi nimi. Vlastní shlukování pak probíhá opakováním kroku sloučení dvou prvků matice s nejmenší vzdáleností do shluku a výpočtem nových vzdáleností mezi tímto shlukem a ostatními prvky.

Postup opakujeme, dokud nevznikne jeden shluk obsahující všechny původní objekty. Existuje více možností výpočtu nových vzdáleností.

- UPGMA (Unweighted Pair-Group Method using Arithmetic averages, metoda průměrné vzdálenosti) je nejpoužívanější metodou [12] a k výpočtu nové vzdálenosti využívá aritmetického průměru původních vzdáleností mezi právě sloučenými prvky (popřípadě již shluky) a prvkem, od něhož novou vzdálenost počítáme. Mějme objekty a, b, c. Jsou-li e_{ab} , e_{ac} a e_{bc} jejich vzdálenosti a e_{ab} je nejmenší, pak novou vzdálenost shluku (ab) od c určíme jako

$$e_{(ab)c} = \frac{(e_{ac} + e_{bc})}{2}. [13] \quad (29)$$

- SLINK (Single Linkage, metoda nejbližšího souseda) novou vzdálenost určí vybráním minimální vzdálenosti mezi právě sloučenými prvky a daným prvkem.

$$e_{(ab)c} = \min(e_{ac}; e_{bc}). [13] \quad (30)$$

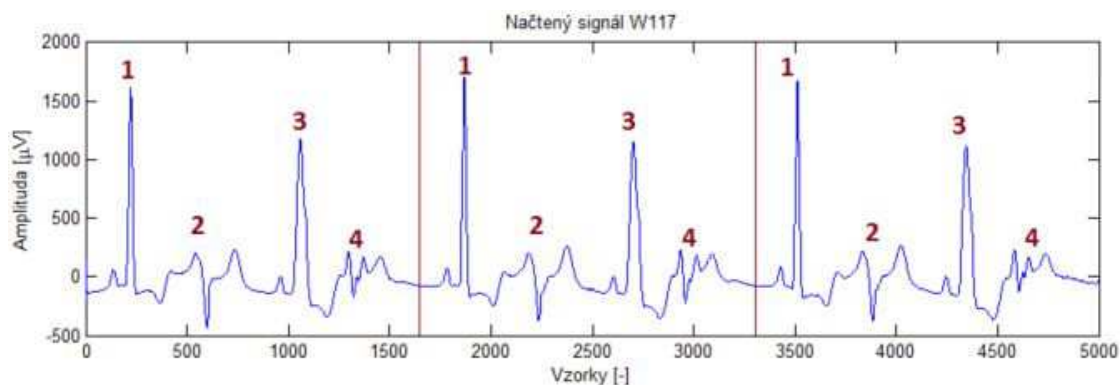
- CLINK (Complete Linkage, metoda nejvzdálenějšího souseda) bere jako novou vzdálenost maximum původních hodnot.

$$e_{(ab)c} = \max(e_{ac}; e_{bc}). [13] \quad (31)$$

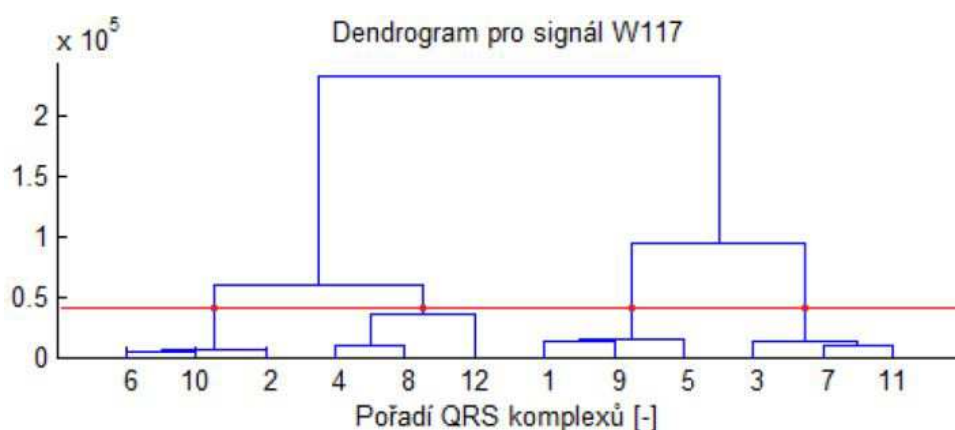
Představené metody jsou pouze výběrem z možností výpočtů nových vzdáleností. Obdobných metod samozřejmě existuje velké množství [13].

Na základě jednotlivých vzdáleností mezi shluky lze vytvořit diagram popisující výsledky analýzy zvaný *dendrogram*. Na jeho vodorovné ose jsou vyneseny všechny cykly, na ose svislé pak jejich vzdálenosti. Vhodně umístěným prahem je možné určit, jak velké shluky požadujeme, to znamená, jak vzdálené ještě mohou být dva prvky zařazené do jednoho shluku.

Problematikou shlukové analýzy EKG pomocí metody DTW se mimo jiné zabývá článek [10]. Příklady klasifikovaného signálu a výsledného dendrogramu jsou v následujících obrázcích (Obrázek 16 a Obrázek 17).



Obrázek 16 Příklad klasifikovaného signálu [10]

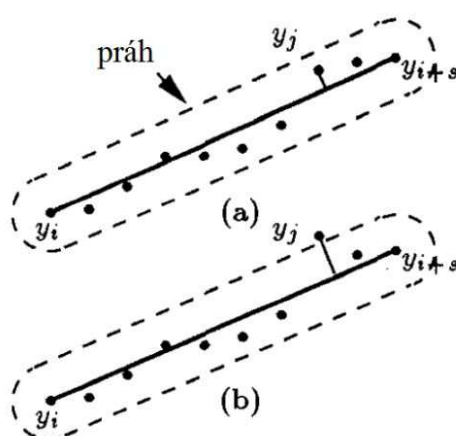


Obrázek 17 Dendrogram příslušný klasifikovanému signálu signálu [10].

3.2 Automatizovaná segmentace EKG

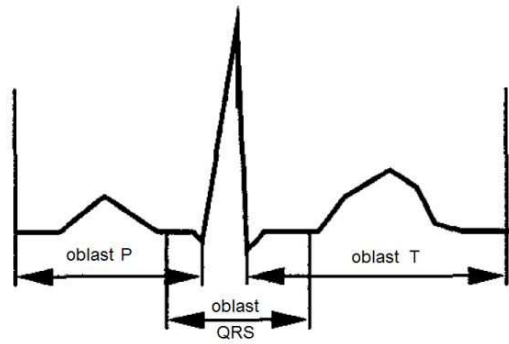
Tato metoda popsaná v [3] sestává z následujících kroků. Prvním je předzpracování tak, jak je uvedeno v kapitole 1.2. Následně aproximujeme signál čarami, jejichž koncové body jsou poté považovány potenciální uzlové body. Segmentace pak závisí na výběru správných bodů. Před aproximací je doporučeno normalizovat cykly na stejný rozsah [3].

Jednotlivé cykly aproximujeme metodou po částech lineární aproximací označovanou PLA (piecewise linear approximation). Mějme signál y délky n a délku kroku metody s . Spojíme body y_i a y_{i+s} a vypočítáme vzdálenosti mezilehlých bodů k úsečce spojující tyto body. Jsou-li všechny tyto vzdálenosti pod úrovní nastaveného prahu, pokračujeme dalším úsekem. V opačném případě spojíme výchozí bod s bodem s největší vzdáleností. Toto zkracování úsečky provádíme, dokud není vzdálenost mezilehlých bodů od úsečky menší než práh. Obrázek 18 představuje myšlenku aproximace. Hodnota prahu je v práci [3] zvolena na 0.01 mV s délkou kroku 32 ms.



Obrázek 18 PLA: (a) aproximace v pořádku, (b) bod je za hranicí, úsečka musí být zkrácena [3].

Dále je cyklus rozdělen na tři překrývající se části. Část před QRS (končící jednu úsečku před R-vlnou), úsek okolo QRS (100 ms) a část za QRS (začínající jednu úsečku za R-vlnou). Viz Obrázek 19.



Obrázek 19 Příklad rozložení jednoho aproximovaného cyklu EKG [3].

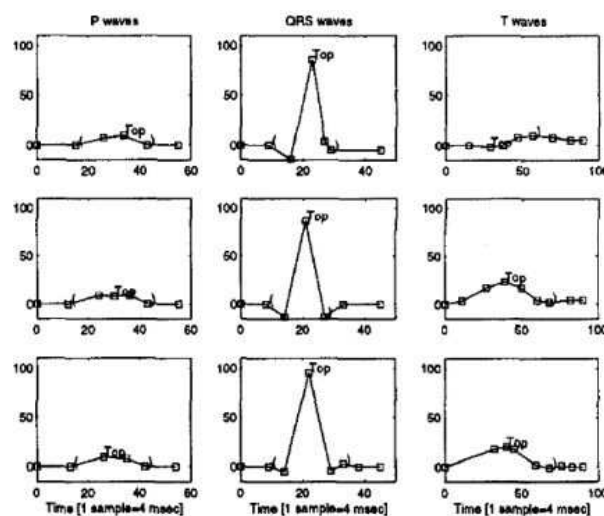
Po těchto úpravách signálu přichází řada na DTW. Při porovnávání dvou cyklů uvažujeme jako vzdálenost pouze absolutní rozdíl sklonu úseček. Zde se pro výpočet používá mírně upravená rovnice kroku (a):

$$\gamma(i, j) = \min \begin{cases} \gamma(i-1, j) & + \delta(i, j) t(a_i) \\ \gamma(i-1, j-1) & + \delta(i, j) (t(a_i) + t(b_j)) \\ \gamma(i, j-1) & + \delta(i, j) t(b_j) \end{cases}, \quad (32)$$

kde $t(a_i)$ a $t(b_j)$ odpovídá trvání aproximační úsečky příslušného cyklu.

Přestože je obvykle naším zájmem celková vzdálenost, v tomto případě velmi přínosná i výsledná cesta. Máme-li informaci o poloze uzlových bodů jednoho cyklu, je jednoduše možné tyto body po zarovnání přenést na cyklus druhý. Otázkou však zůstává, jak vybrat dobrý referenční cyklus.

Z důvodu značné proměnlivosti mezi jednotlivými cykly, ať už v rámci jednoho záznamu či mezi záznamy různými, by bylo zapotřebí velké databáze zachycující rozsah všech možných průběhů. Je možné vytvořit menší počet reprezentativních úseků P, QRS a T (v řádu jednotek) a jejich vzájemnou kombinací pokrýt větší počet možností (Obrázek 20).



Obrázek 20 Příklady referenčních vln, jejichž kombinací lze vytvořit 27 různých průběhů [3].

Práce [8] nabízí v podstatě stejné řešení problému a je částečně rozšířením popsané metody. Rozdílem je využití aproximační metody APCA (adaptivní po částech konstantní aproximace) a metody DDTW.

3.3 Příklady jiných využití

Vzhledem k faktu, že většina metod využívajících DTW se často liší pouze mírně odlišným předzpracováním, jinak se jedná o porovnávání s referenčním vzorem, uvádím dále pouze příklady dalších možných využití bez bližšího rozboru.

Klasifikace EKG cyklů uvedená ve [4] je založena na porovnání minimální střední hodnoty rezidua vzniklého při borcení časové osy. Spadá-li hodnota do známého rozsahu, je vstupní cyklus přiřazen příslušné třídě. Pokud ne, je vytvořena třída nová.

[5] se zabývá využitím metody **DTW pro biometrickou identifikaci založenou na snímání EKG**. Přestože se záznam EKG jako biometrická identifikační metoda pravděpodobně nikdy neprosadí vzhledem k vysoké variabilitě EKG (názor autora), je dalším zajímavým využitím algoritmu. V předzpracování jsou vybrány reprezentativní vzorky, které jsou následně aproximovány metodou zvanou PLR (shodná s PLA) za účelem redukce množství dat. Cykly jsou pak porovnávány s uloženými vzory, kde kritériem je nejmenší vzdálenost DTW, která musí být současně nižší než empiricky zvolený práh.

4 Programové realizace algoritmů

4.1 Použité signály

Pro účely praktické části bude využito databáze Common Standards for quantitative Electrocardiography (CSE). Jde o standard pro návrhy a testování měřících a vyhodnocovacích algoritmů v elektrokardiografii. Zaznamenané signály s vzorkovací frekvencí 500 Hz obsahují 5000 vzorků, což odpovídá 10 vteřinám. V některých případech je diagnosticky významných pouze 8 vteřin záznamu.

Signály jsou rozděleny do několika datových sad, dále pak podle původu na originální a umělé a podle počtu svodů na Frankovy třísvodové (X, Y, Z) a klasické 12svodové [18].

V rámci vývoje byly jak za pomoci nezávislých odborníků, tak pomocí specializovaných programů signály rozměřeny a klasifikovány. Výsledky dostupné v souboru MRESULTS.xls byly převedeny do matice v souboru .mat. Byly odstraněny nadbytečné informace o dílčích výsledcích klasifikací jednotlivých programů. Řádek matice odpovídá signálu, sloupce pak postupně začátku referenčního cyklu, jeho konci, začátkům a koncům vlny P, komplexu QRS a vlny T.

Výsledky jsou dostupné pouze pro signály označené MO1, proto budou právě ty využity přednostně. Konkrétně MO1_001_12 až MO1_125_12 a MO1_001_3 až MO1_125_3. Klíčové body jsou bohužel vždy určeny pouze u jednoho - referenčního - cyklu záznamu, tato nevýhoda je však kompenzována vyšším počtem signálů. Pro některé signály nebyly nebo nemohly být referenční hodnoty stanoveny, příslušná buňka tabulky pak obsahuje nulu a tyto signály nebudou dále využity. Vyřazeny byly také signály značně poškozené či s atypickým průběhem.

Pro testování algoritmů zůstává k dispozici 97 signálů. Předpokládáme-li, že každý z 12 svodů se alespoň částečně liší průběhem, mohou být v celkovém součtu metody borcení časové osy testovány až na 1164 průbězích referenčních cyklů signálu EKG.

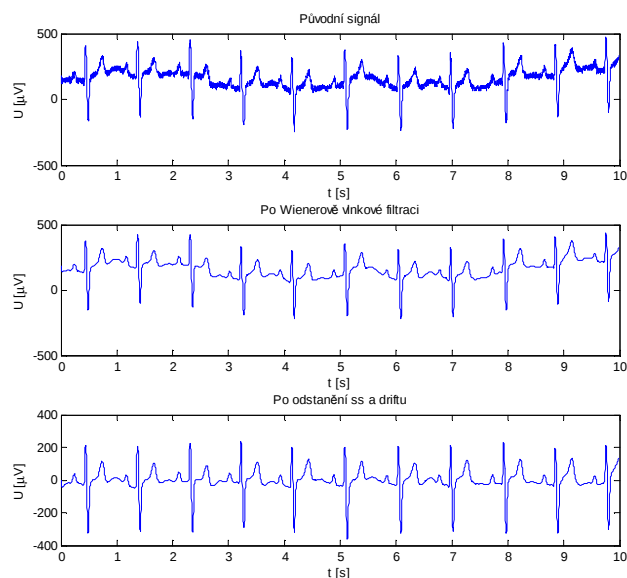
4.1.1 Filtrace

Kolísání nulové izolinie a stejnosměrná složka signálu byly potlačeny metodou nulování spektrálních čar horní propustí s mezní frekvencí 0,6 Hz. Metoda byla i využita i při potlačení síťového rušení, zde pásmovou zádrží s mezními frekvencemi 49,9 a 50,1 Hz. Hranice byly zvoleny s ohledem na požadavky filtrace a na vzdálenost spektrálních čar při dané vzorkovací frekvenci a počtu vzorků. Vzdálenost čar činí 0,1 Hz, čili ani nebylo možné nastavení v jemnějším rozsahu. Filtraci provádí funkce *nulovani* (Dodatek 2).

Problémem některých signálů je značné rušení myopotenciály. Předchozí metoda je zde v zájmu zachování informačního obsahu elektrokardiografického signálu nepoužitelná. Zde je nezbytné použít sofistikovanější metody. Na základě práce [16] byla vybrána Wienerova

vlnková filtrace a aplikován algoritmus v rámci práce vytvořený. Filtraci provádí funkce *Vlnkovy_filtr*.

Na obrázku (Obrázek 21) je zřejmé jak vymizení rušení myopotenciály, tak ustálení signálu na nulové hladině. Odstranění driftu není vždy naprosto ideální, například, je-li kolísání příliš veliké, jako u svodu V1 signálu MO1_003_12.

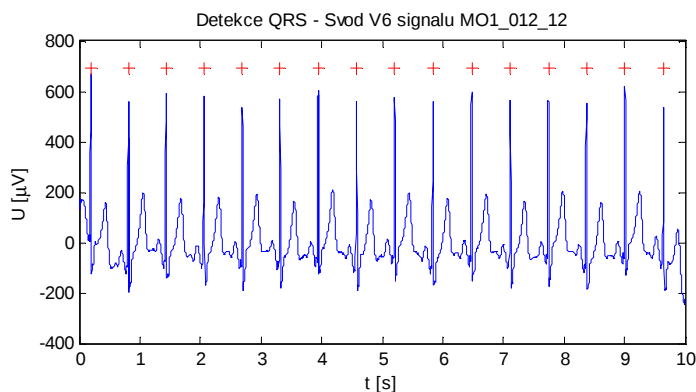


Obrázek 21 Svod I záznamu MO1_001_12 a odstranění rušení.

Za účelem vyvarování se neustálého opakování časově relativně náročných filtrací, byly všechny dostupné signály v počátcích práce přefiltrovány a uloženy s předponou „f“. Například fMO1_025_12.

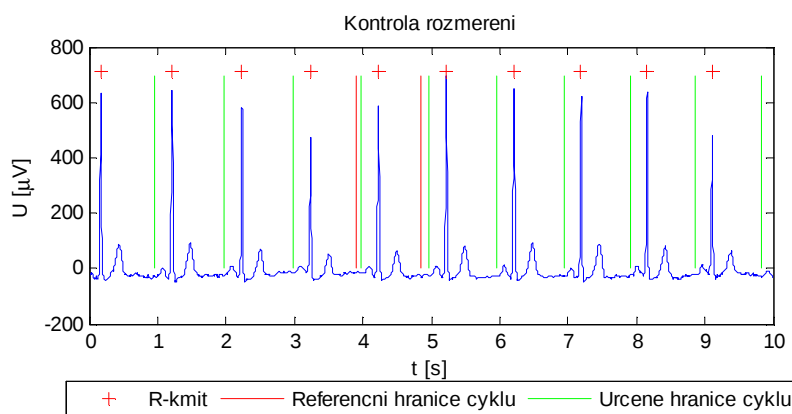
4.1.2 Rozměření

K detekci polohy R vlny byl převzat detektor navržený v [9] založený na vlnkové transformaci s mateřskou vlnkou *bior1.5* v měřítku 15. Příklad detekce je na obrázku níže (Obrázek 1).



Obrázek 22 Detekce QRS

U všech signálů bylo provedeno rozměření na základě metody popsané v kapitole 1.2.2 U krajních cyklů je hodnota začátku (konce) stanovena na základě následujících (předchozích) hodnot. Je-li hranice „mimo“ délku signálu, není tento cyklus brán v potaz. Stanovené hranice cyklů byly porovnány s referencí a kontrolovány i vizuálně. Kontrola prokázala nevhodnost použití rozměření v poměru 75/25 %. Příklad kontroly pro svod aVL signálu MO1_066_12 je na obrázku (Obrázek 23). Hranice cyklů jsou vzdálené referenci a velmi blízko vlně P.



Obrázek 23 Kontrola rozměřovacího algoritmu

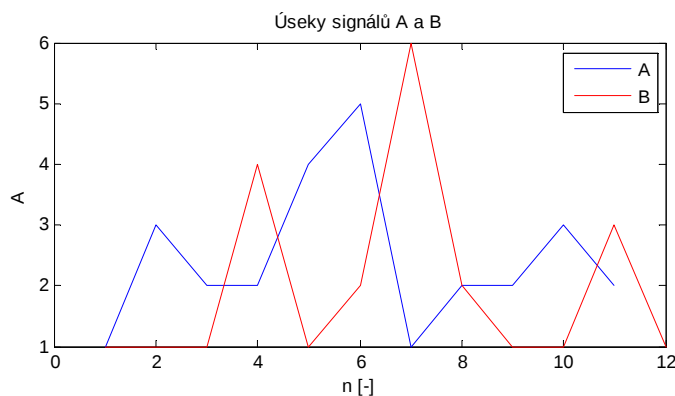
Empiricky byl pro tuto sadu signálů stanoven poměr rozměření 70/30 % poskytující uspokojivější výsledky.

4.2 DTW

Samotné borcení časové osy provádí funkce *dtw*. Její algoritmus bude pro názornost přiblížen na kratších úsecích signálů, následně při porovnávání cyklů EKG. Mějme vektory

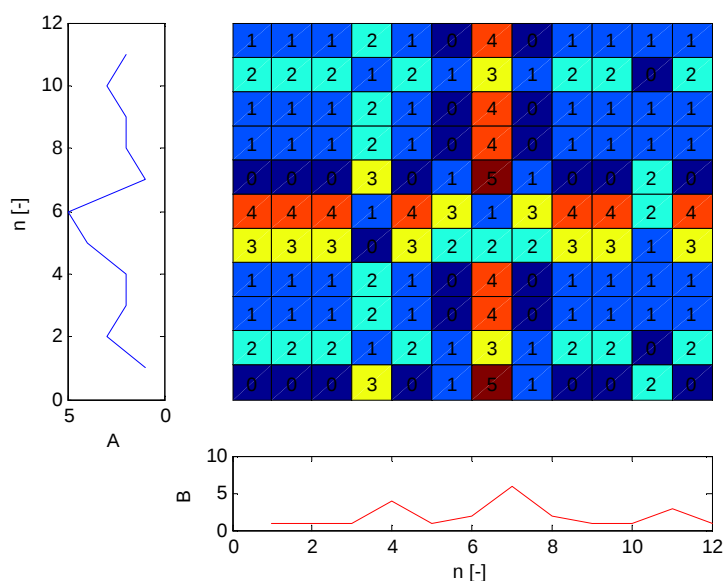
$$\begin{aligned} A &= \{1, 3, 2, 2, 4, 5, 1, 2, 2, 3, 2\} \quad \text{a} \\ B &= \{1, 1, 1, 4, 1, 2, 6, 2, 1, 1, 3, 1\}. \end{aligned} \quad (33)$$

Ty si vlastně lze představit jako kratší úseky signálů.



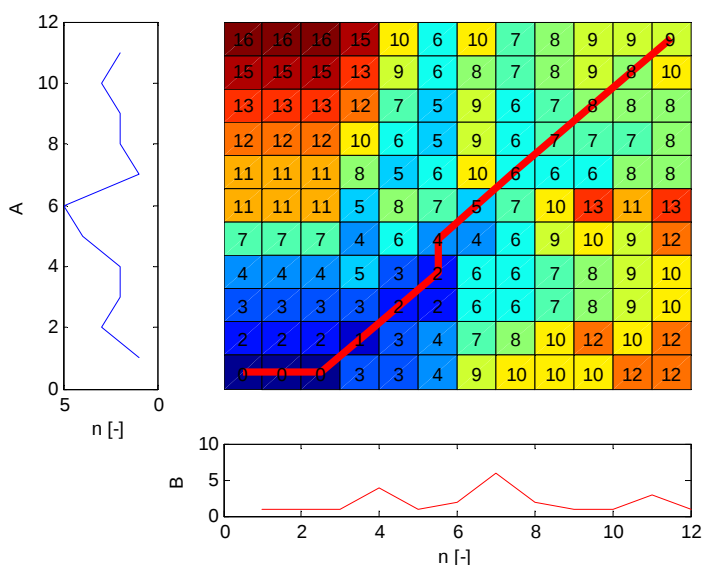
Obrázek 24 Úseky porovnávaných signálů

Na základě rovnice (17) sestojíme matici lokálních vzdáleností rozměru 11×12.



Obrázek 25 Matice lokálních vzdáleností

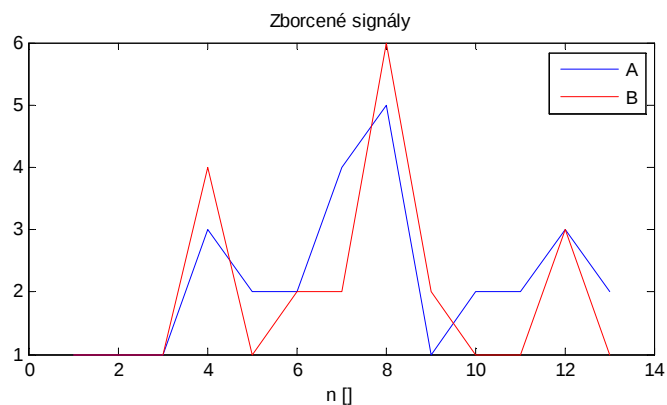
Matici lokálních vzdáleností následně přepočítáme na matici kumulovaných vzdáleností s využitím kroku (a). V pravém horním rohu je nyní vzdálenost DTW obou signálů. Z této pozice je potom trasována výsledná cesta, která je také v Obrázek 26 vyznačena.



Obrázek 26 Matice kumulovaných vzdáleností s vyznačenou cestou.

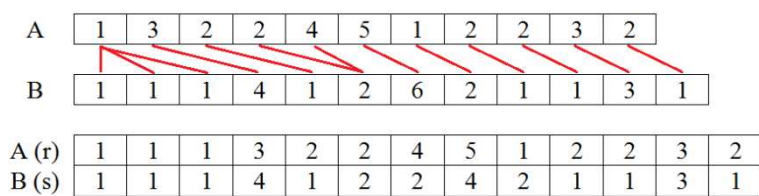
Interpretace bortivé cesty je následující. Vyjdeme-li z levého dolního rohu, první dva kroky započítáváme první při vzorky z B, zatímco „držíme“ hodnotu prvního vzorku z A. V dalších třech krocích plyne čas normálně. Poté jeden krok držíme hodnotu z B, zatímco přidáme další z A. Po zbytek cesty již čas plyne opět normálně.

Na signálech se zborcenou časovou osou je při porovnání s původními signály patrné zarovnání výrazných vrcholů (Obrázek 27).



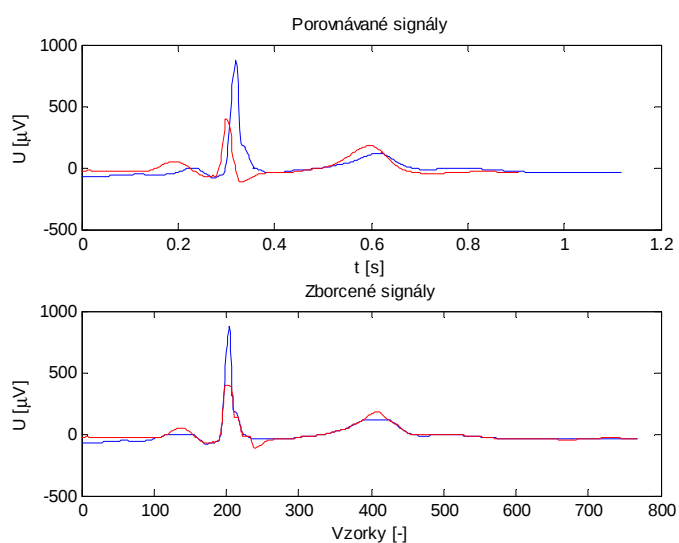
Obrázek 27 Signály po borcení časové osy

Názorně je principi a výstup předchozího borcení vidět při zobrazení na vlastních posloupnostech A a B (Obrázek 28).

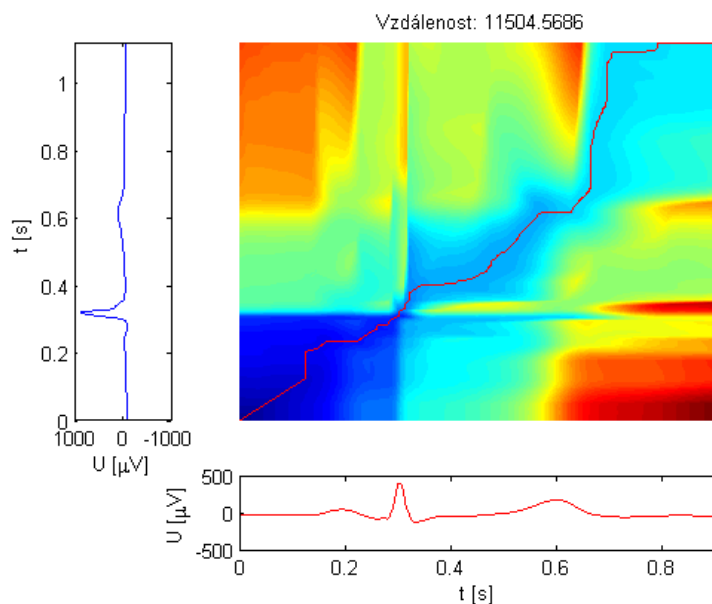


Obrázek 28 DTW s čísly

Pro příklad zarovnání dvou cyklů EKG signálu byly náhodně vybrány signály MO1_025_12 a MO1_106_12. Porovnávány budou referenční cykly svody I.



Obrázek 29 DTW pro dva cykly EKG



Obrázek 30 Výsledná cesta při zarovnávání vybraných cyklů EKG.

4.2.1 Vliv rozdílů v porovnávaných signálech

Doposud bylo využito algoritmu DTW v nejzákladnější formě, bez omezení kladených na vyhledávací algoritmus. Jediná omezení implementovaná v algoritmu byla s ohledem na průběh výsledné cesty (okrajové podmínky, monotónnost, spojitost). Úkolem následujících kapitol je pochopit chování algoritmu při aplikaci na jednotlivé cykly EKG a na základě získaných výsledků optimalizovat kladená omezení za účelem docílení co nejlépejších výsledků.

Jsou-li porovnány dva naprosto totožné cykly, výsledek je nasnadě. Vzdálenost je nulová a cestu tvoří úhlopříčka matice vzdáleností. Předpokladem pro další testování je výskyt průběhem totožných cyklů v EKG záznamu. Jediným rozdílem bude délka, nulová úroveň, přítomnost změn nebo použitý krok. Testován bude vždy pouze referenční cyklus svodu I u všech signálů. Bude použit krok podle rovnice 11.

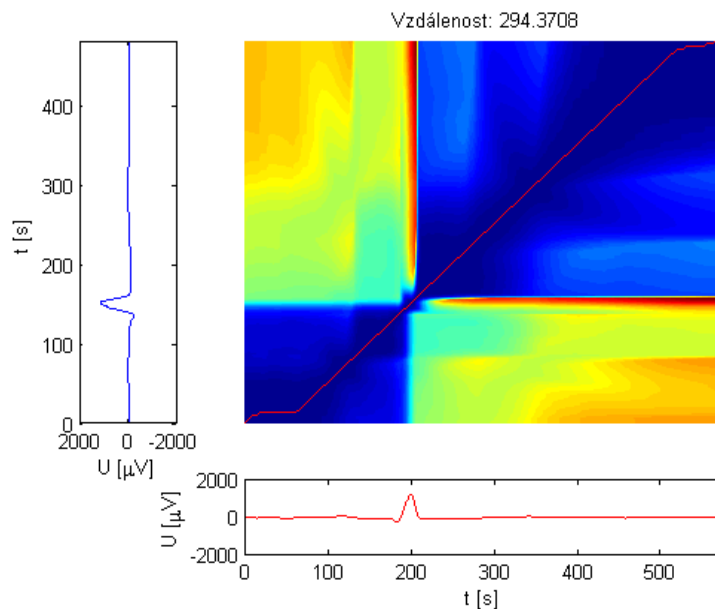
Vliv délky porovnávaných úseků

Tento pokus je navržen k pochopení reakce algoritmu na nedokonalé rozměření signálu. Změna mezi cykly bude taková, že u druhého cyklu budou změněny hranice v rozmezí $\pm 10\%$ délky cyklu. Změna hranice je navržena empiricky tak, aby při provedené změně částečně zasahovala do vln cyklu. Obdržíme 8 porovnání změněné délky pro každý cyklus. Výsledná cesta bude logicky opět téměř úhlopříčná. Pouze v rozích matice se algoritmus musí vypořádat se změnami. Závěry podkapitoly pomohou jak s nastavením prahu pro určení podobnosti, tak možnosti nastavení globálních omezení v rozích matice (Obrázek 13).

Výsledky jsou uloženy v souboru *delka.mat*. Bližší popis následuje v Dodatek 3. Vzdálenost cyklů se při vyloučení extrémních hodnot u signálu MO1_023_12 pohybovala v rozmezí $\langle 26; 3963 \rangle$. Většina hodnot (82 %) se pohybovala do vzdálenosti 1000. Tato

informace je však pouze informativní, neboť se vzdálenosti mohou značně lišit v závislosti na podmínkách výpočtu.

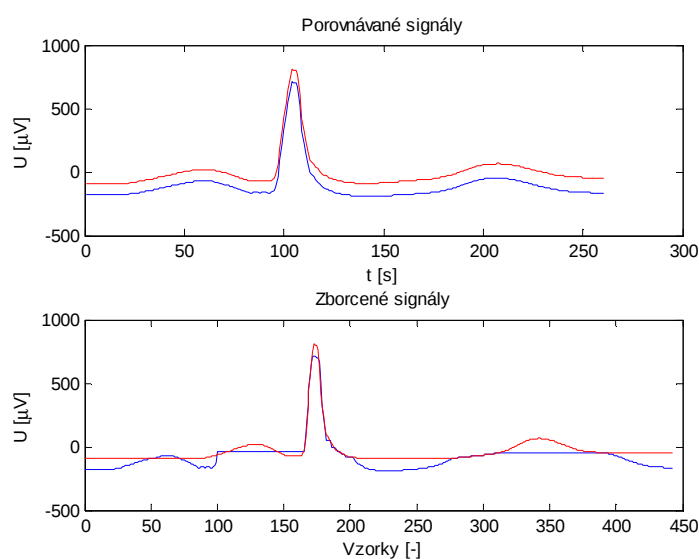
Maximální počet kroků v horizontálním nebo vertikálním směru by měl poukázat na nežádoucí dilataci úseků. Pohyboval se v rozmezí $<1; 87>$ s průměrem 34,32 (medián = 36).



Obrázek 31 DTW při různém rozměření téhož cyklu.

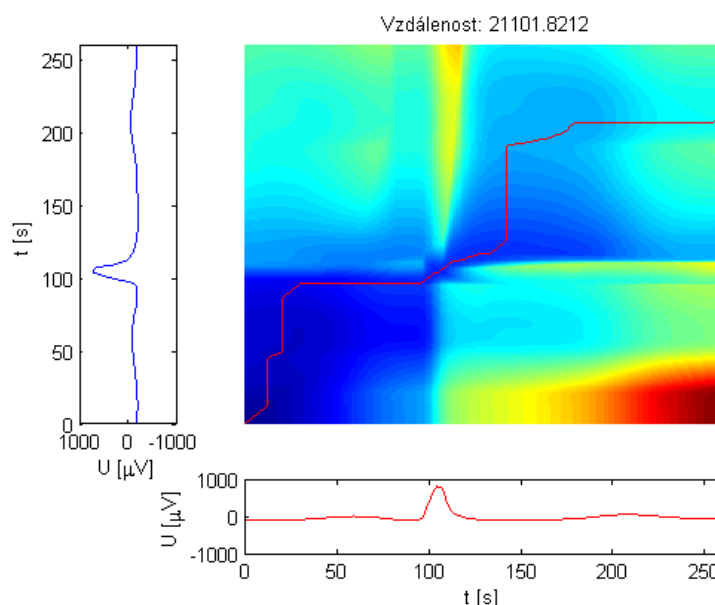
Vliv rozdílné úrovně

Při porovnávání shodného cyklu hraje roli i úroveň nulové linie signálu. Při porovnávání filtrovaného signálu a signálu postiženého driftem může dojít ke špatné interpretaci hodnot a tím nesmyslné dilataci některých úseků.



Obrázek 32 Nesprávné zarovnání signálů.

U referenčního cyklu signálu MO1_115_12 je dilatace pozorovatelná hned v několika případech (Obrázek 32). Na bortivé cestě se dilatace projevují jako mnoho kroků v horizontálním či vertikálním směru, jak bylo již zmíněno výše.



Obrázek 33 Bortivá cesta pro referenční cyklus signálu MO1_115_12.

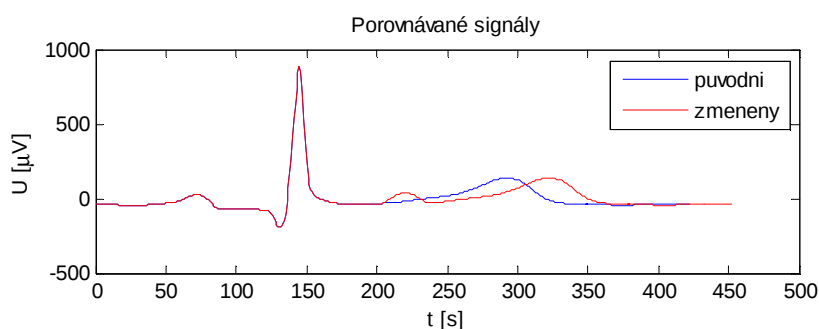
Výsledky jsou uloženy v souboru *uroven.mat*. Bližší popis následuje v Dodatek 3.

Vzdálenost cyklů se pohybovala v rozmezí $<2\ 088; 550\ 140>$. Většina hodnot (87 %) se pohybovala do vzdálenosti 100 000. Hodnoty závisí na rozdílu úrovní porovnávaných cyklů.

Počet kroků v horizontálním nebo vertikálním směru se pohyboval v rozmezí $<0; 329>$ s průměrnou hodnotou 78, 52 (medián = 71, 5).

Vliv přítomnosti změn

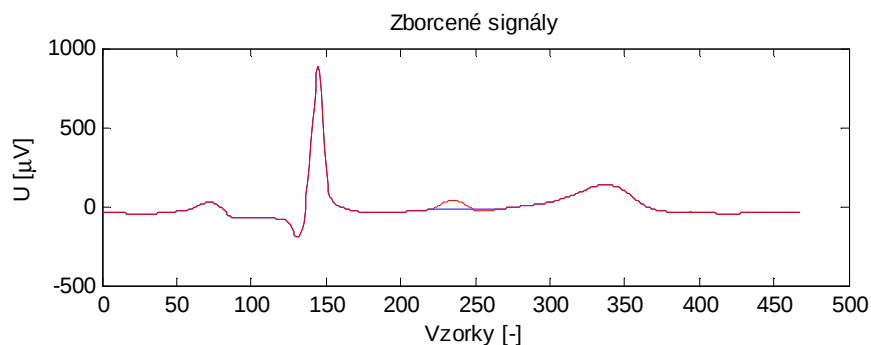
V tomto pokusu je do náhodné pozice v signálu vložena vlna délky 30 vzorků (~ 60 ms) velikosti $70\ \mu\text{V}$. Výsledky jsou uloženy v souboru *zmena.mat*.



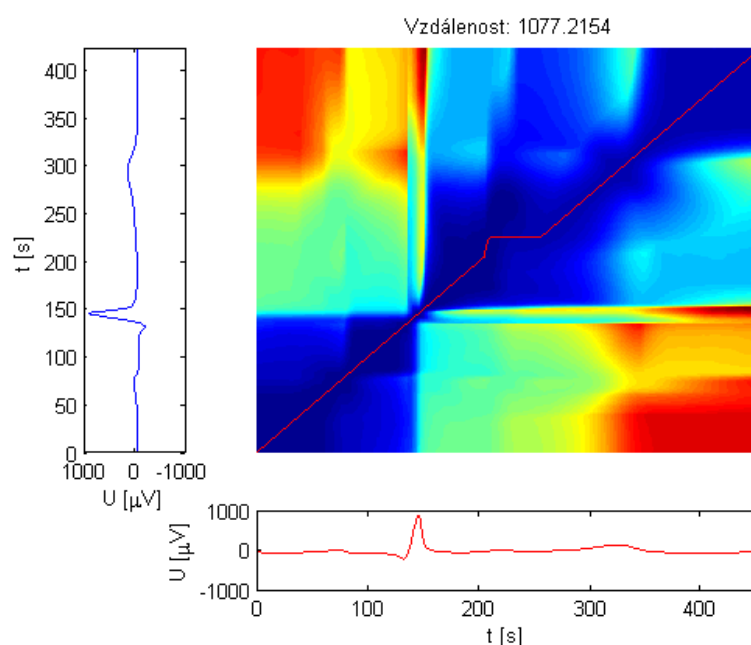
Obrázek 34 Změna v signálu.

Algoritmus ve většině případů změnu bez větších problémů „obešel“ dilatací úseku původního cyklu. Porovnání signálu MO1_038_12 a cesta je dále na Obrázek 35 a Obrázek 36.

Výsledná vzdálenost se logicky odvíjela od úrovně a polohy změny. Byla v rozsahu $\langle 382; 1335 \rangle$. Počet kroků ve vertikálním nebo horizontálním směru byl maximálně 56 a průměrem 17,61 (medián = 16,5).



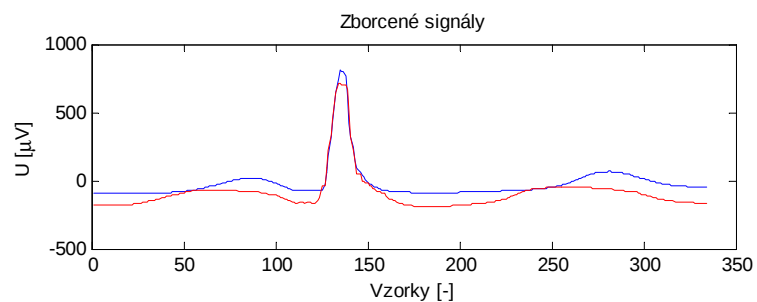
Obrázek 35 Zarovnané signály,



Obrázek 36 Bortivá cesta.

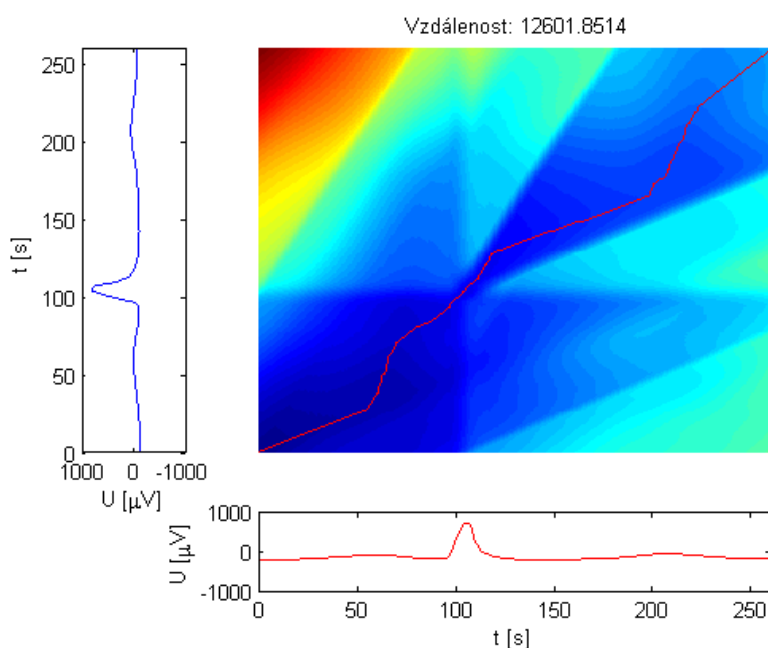
Vliv použitého kroku

Využití ostatních kroků popsaných v kapitole 2.1.2 při testování změny úrovně a přítomnosti změny vedlo k výrazně odlišným výsledkům. Při využití kroku podle rovnice 13 nedochází při různé úrovni signálu k tolik výrazným dilatacím. Výsledky uložené v *uroven_k2.mat* (Dodatek 3) prokazují mnohonásobně menší počet pohybů jak v horizontálním, tak vertikálním směru. Rozdíl při zarovnání je patrný při porovnání Obrázek 37 a Obrázek 32. Jedná se o totožný signál.



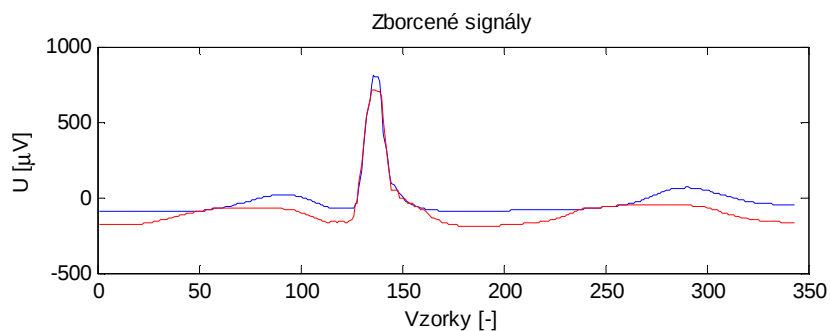
Obrázek 37 Zarovnání signálu MO1_115_12 s krokem podle rovnice 13

Porovnáním vyznačených cest Obrázek 33 a Obrázek 38 je opět znatelný markantní rozdíl.

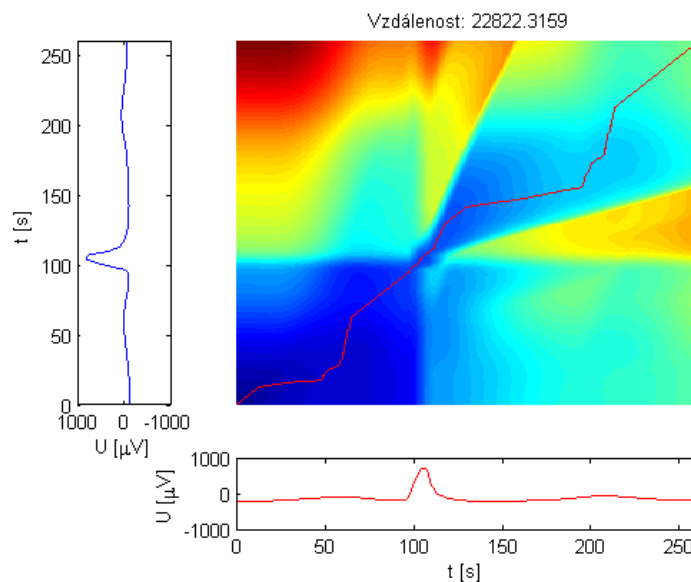


Obrázek 38 Bortivá cesta při využití kroku (b).

Využitím kroku (c) dosáhneme mírně lepších, ve většině případů srovnatelných výsledků jako v předchozím případě, ovšem s vyšší výpočetní náročností.

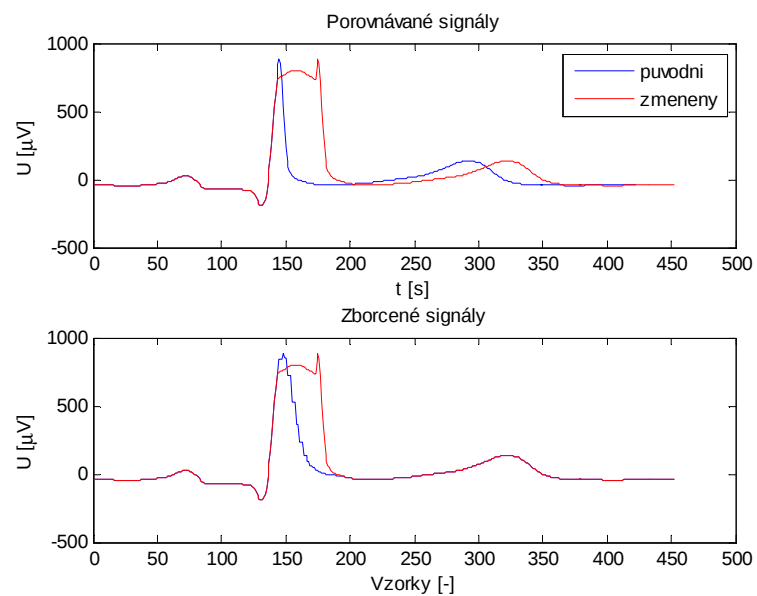


Obrázek 39 Zarovnání při kroku (c).

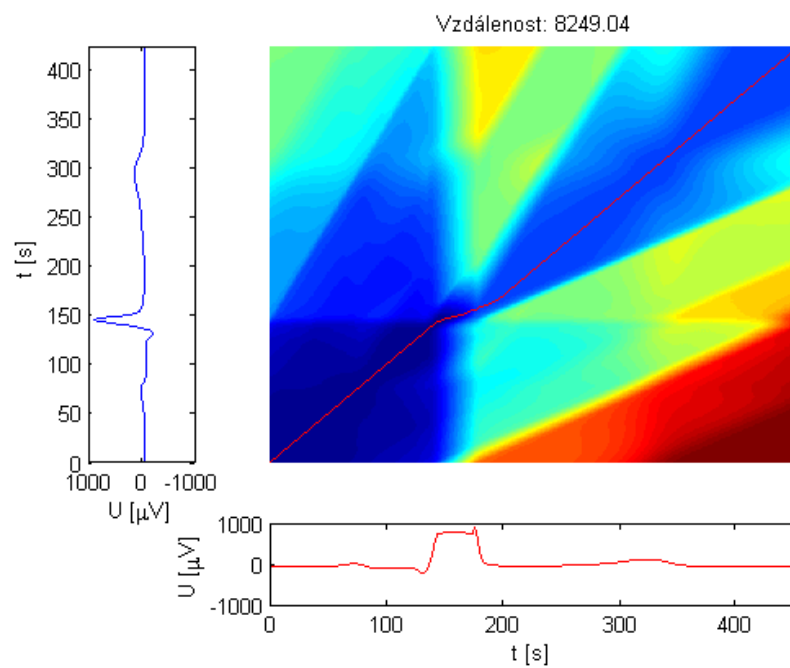


Obrázek 40 Bortivá cesta při kroku (c).

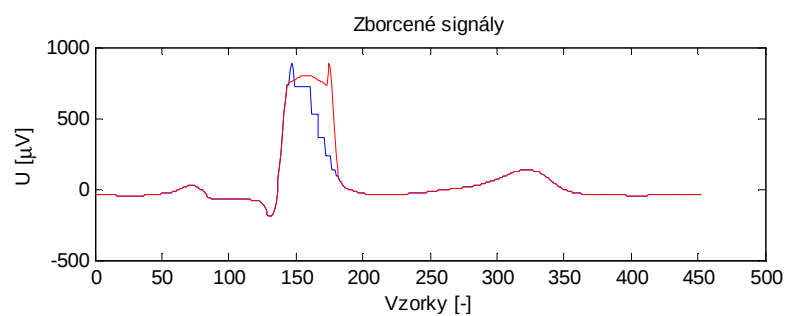
Vyskytuje-se v signálu náhodná změna, ukazuje se výhodnějším využití jiných kroků než (a). Pro případ (b) jsou výsledky uloženy v *zmena_k2.mat*. Hodnocení i pro odpovídající signály může být však zavádějící, neboť byla pozice změny vybírána náhodně. Nicméně už jen vizuální kontrolou všech srovnání jsou zřejmé výhody tohoto kroku. I přes rozdíl nedochází k výrazné dilataci jedné hodnoty, jako tomu bylo v případech kroku (a). I cesta vykazuje „hladký“ průběh (Obrázek 42).



Obrázek 41 Zarovnání signálů.



Obrázek 42 Výsledná cesta.



Obrázek 43 Reakce algoritmu na změnu při využití kroku (c)

Krok podle (c) opět přináší lepší výsledky (*zmena_k3.mat*) než (a), v porovnání s (b) však v tomto případě zaostává. Algoritmus má při snaze vyvážit rozdíl tendence k vytváření „schodů“ (Obrázek 23 Kontrola rozměrovacího algoritmu). V příkladu na obrázku byla úmyslně zvolena totožná změna a signál jako v předchozím případě.

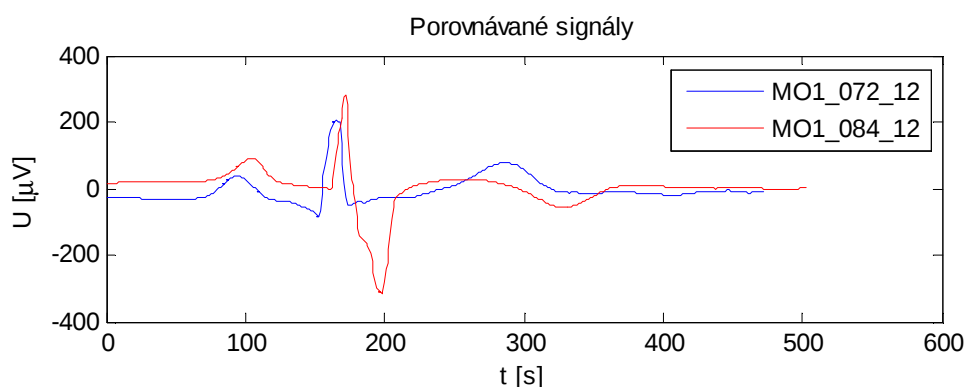
4.2.2 Omezení

Provedené pokusy prokázaly, že nebude-li využito „základního“ kroku (a), nebude nezbytně nutné využívat další lokální omezení jako omezení strmosti popsaná v kapitole 2.1.2. Jako dostatečné lokální omezení algoritmu je možné považovat využití kroků (b) nebo (c). Tento závěr má výhodu především díky faktu, že samotné určení maximálního přípustného počtu kroků v jednom směru je nesnadné. Porovnáním dosažených výsledků je možné odhadnout, že zvolená hranice by se měla pohybovat cca do dvaceti vzorků v jednom směru (odhad - netestováno).

S ohledem na nižší výpočetní nároky a srovnatelné výsledky s případy využití kroku (c) bude při dalších aplikacích algoritmu přednostně využíváno kroku (b).

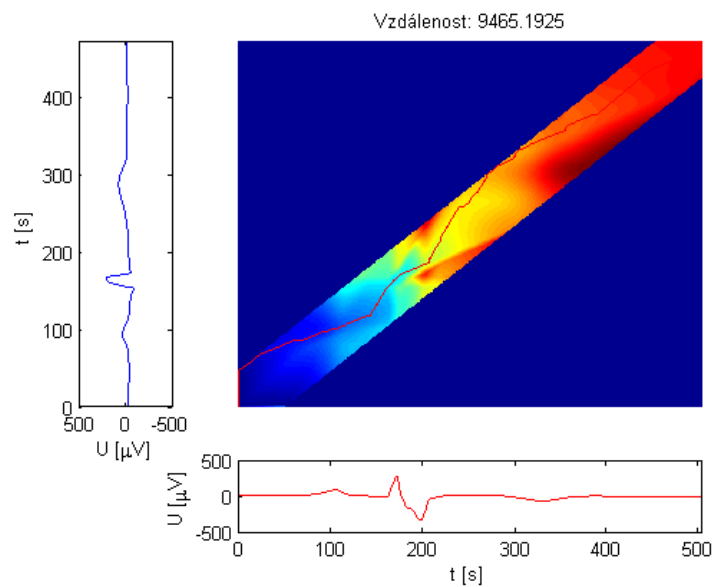
Globální omezení algoritmu, jsou-li zvolena správně, mohou zlepšit přesnost klasifikace, směřovat cestu do patřičných mezí a zvýšit rychlost výpočtu vzdálenosti[20]. Obvykle je hodnota stanovena jako 10% vzdálenost od diagonály matice [22].

Do funkce DTP bylo globální omezení implementováno s možnostmi kroku (a) a (b). Náhodně byl vybrán signál MO1_072_12, který byl porovnán se všemi ostatními s využitím 10% a 20% šířkou okna a s kroky (a) a (b). Největší rozdíl Výsledných vzdáleností byl ve srovnání se signálem MO1_084_12, který byl proto vybrán jako ilustrační případ. Výsledky v *glob_72_10_k1.mat*, *glob_72_20_k1.mat*, *glob_72_10_k2.mat* a *glob_72_20_k2.mat* (**příloha**).

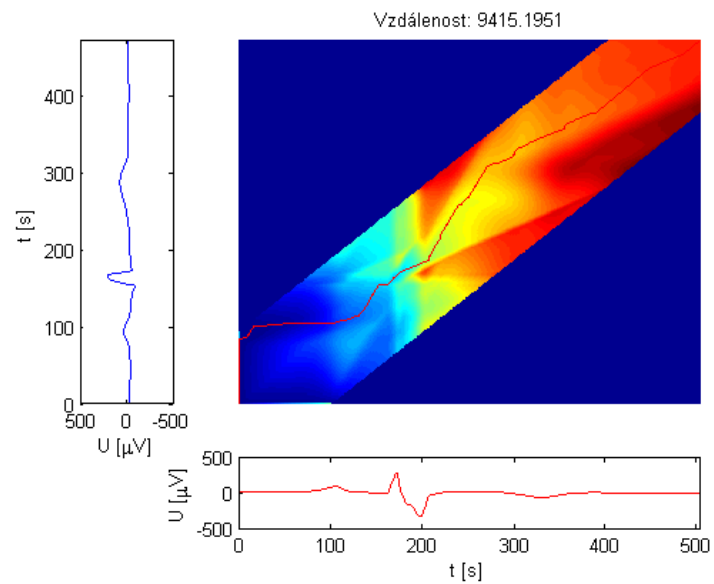


Obrázek 44 Signály s největším rozdílem vzdáleností s ohledem na použitou šířku okna

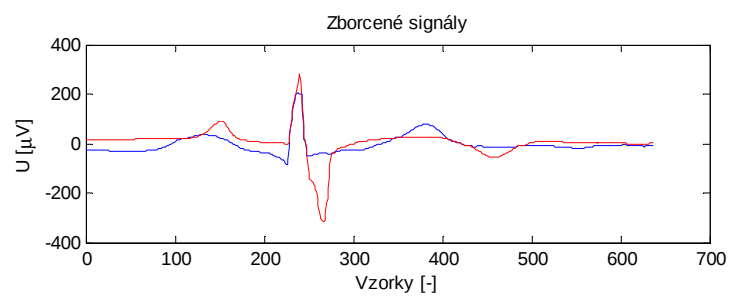
Rozdíly v průběhu bortivé cesty jsou ilustrovány v Obrázek 45 a Obrázek 46. Rozdíl v borcení časové osy signálů pak v Obrázek 47 a Obrázek 48.



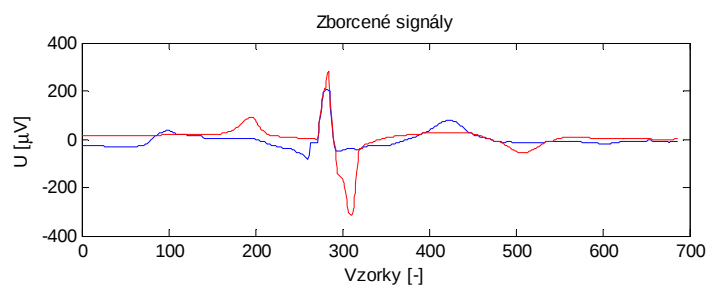
Obrázek 45 Bortivá cesta pro globálním omezení s šířkou okna na 20% šířky diagonály.



Obrázek 46 Bortibá cesta při globálním omezení s šířkou okna na 20% šířky diagonály.



Obrázek 47 Signály se zborcenou časovou osou při globálním šířkou okna na 10% šířky diagonály.



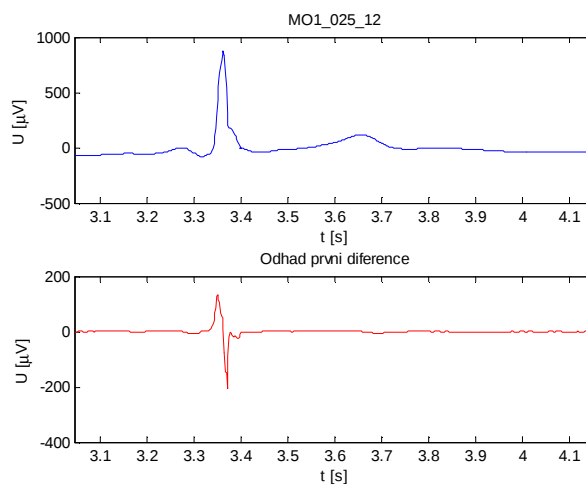
Obrázek 48 Signály se zborcenou časovou osou při globálním šířkou okna na 20% šířky diagonály.

Přestože se vzdálenosti příliš neliší, je při využití užšího okna znatelné příznivější zarovnání. Zároveň je možné potvrdit předpokládanou nižší výpočetní náročnost při využití globálních omezení. Při využití matlabovské funkce `cputime` byl v průměru naměřen o 53,5 % nižší výpočetní čas při porovnávání referenčního cyklu signálu MO1_001_12 a referenčních cyklů signálů ostatních. Výsledky jsou uloženy v *CPU.mat* (Dodatek 3).

Tabulka 1 Příklady rozdílných výpočetních dob

Signál	Čas s omezením [s]	Čas bez omezení [s]	Rozdíl [%]
MO1_004_12	0,2964	0,6708	55,8139
MO1_022_12	0,2964	0,6708	55,8139
MO1_042_12	0,3276	0,6864	52,2727
MO1_051_12	0,2184	0,4368	50
MO1_086_12	0,3120	0,7644	59,1836
MO1_088_12	0,2808	0,6552	57,1428

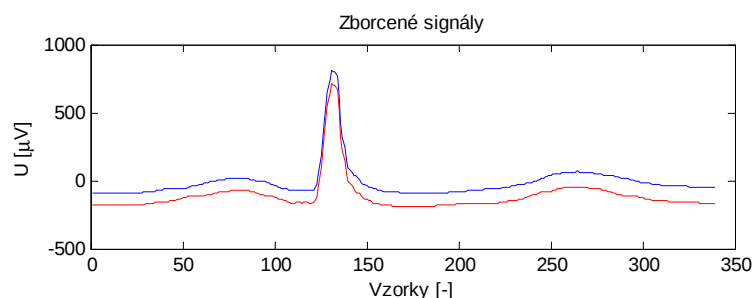
4.2.3 DDTW



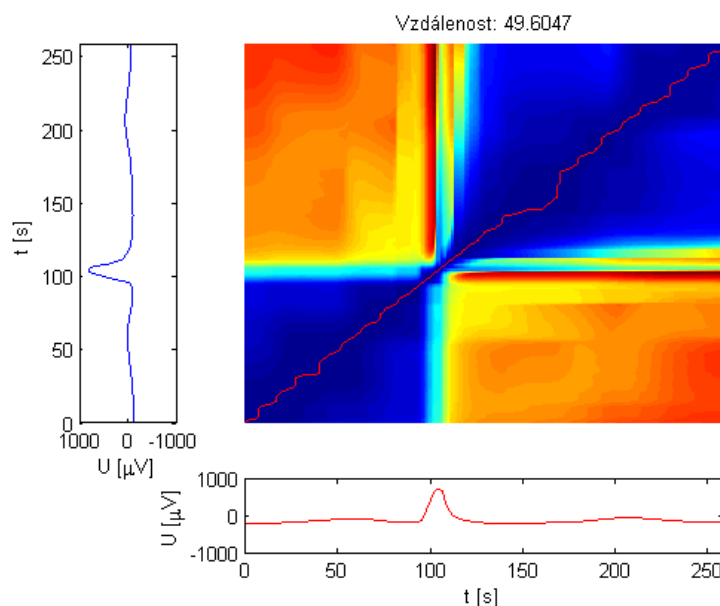
Obrázek 49 Signál a odhad jeho první difference podle 28.

Borcení časové osy s využitím první difference porovnávaných posloupností se od klasické metody liší pouze výpočtem odhadu první difference (Obrázek 49). Výpočet je implementován ve funkci dtw , kde je proveden při daném nastavení proměnné $DDTW$.

První difference je nezávislá na úrovni signálu. V porovnání s klasickou metodou přináší DDTW optimálnější výsledky, zároveň příliš nezáleží na použitém kroku. Pro možnost porovnání (Obrázek 50 a Obrázek 51) je algoritmus opět demonstrován na referenčním cyklu signálu MO1_115_12 s využitím kroku (a).



Obrázek 50 DDTW signálu MO1_115_12.



Obrázek 51 Bortivá cesta při zarovnání signálu MO1_115_12 s různými úrovněmi.

Chování algoritmu při vložení změny do posloupnosti je srovnatelné s DTW.

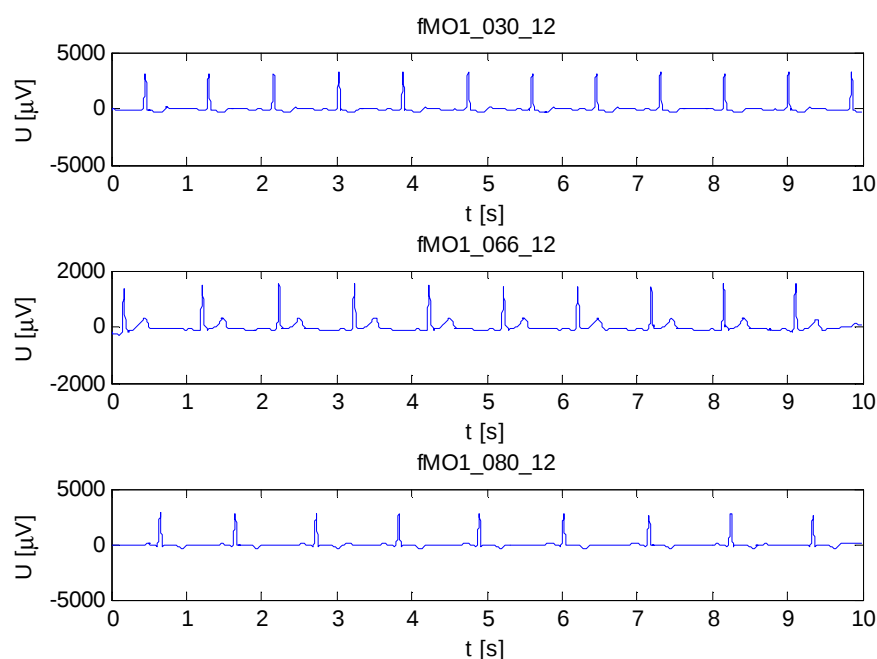
4.2.4 Shrnutí

Z předcházejících kapitol ve znamení snahy o nalezení optimálního nastavení algoritmu vyplívá následující. Při určování vzájemné vzdálenosti cyklů EKG pomocí metody borcení časové osy je výhodné s ohledem na možné změny délky, úrovně a podobně, využít lokálního omezení v podobě kroku (b), s globálním omezením s oknem šířky 10 % diagonály matice kumulovaných vzdáleností a metody borcení využívající odhadu první difference.

4.3 Shluková analýza

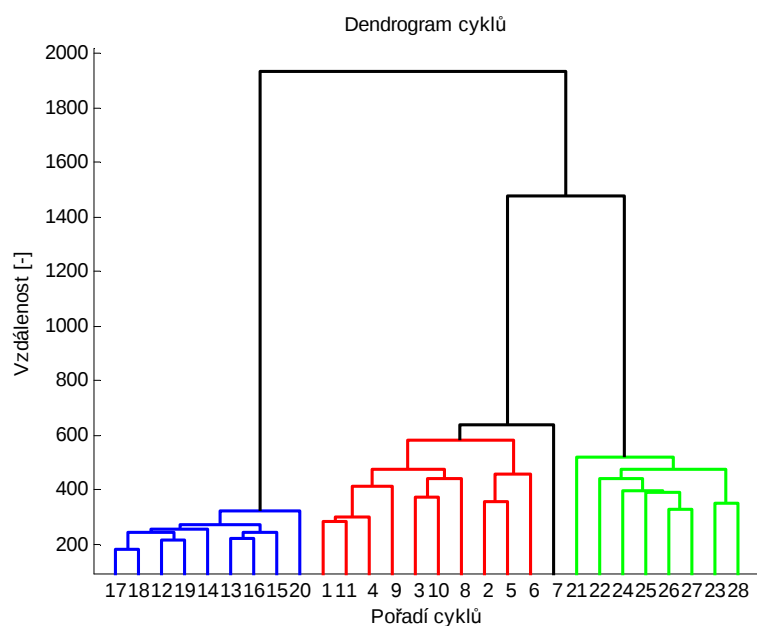
Shluková analýza jako taková je článkem následujícím za určením vzájemných vzdáleností jednotlivých elementů, v tomto případě cyklů. Zmíněna je tu především z důvodu jejího využití při návrzích v další části textu. Analýza je provedena funkcí *shlu_anal.m* (Dodatek 3). Z vybraných signálů (a jejich svodů) jsou extrahovány cykly, ty jsou následně vzájemně porovnány. Shlukování je provedeno pomocí matlabovské funkce *linkage*.

V rámci této funkce je možné využít mnoha metod shlukování, jelikož není tato práce zaměřena na jejich porovnávání, byla využita metoda UPGMA. Pro nastavení DTW platí závěry ve shrnutí předchozí kapitoly. Porovnávány budou pro ilustraci svody V5 náhodně vybraných signálů MO1_030_12, MO1_065_12 a MO1_080_12.



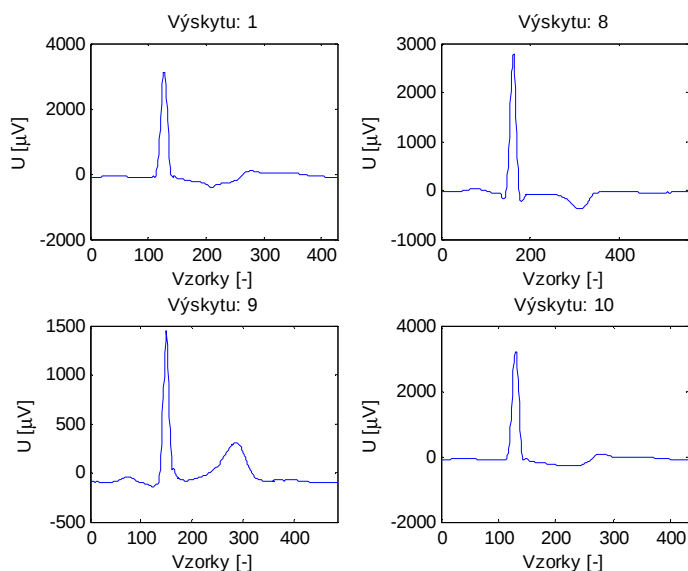
Obrázek 52 Signály vybrané pro shlukovou analýzu

Vzniklá matice vzdáleností má mít rozměr 28×28 , její naplnění je proto časově poměrně náročné, i když je symetrická a stačí vyplnit polovinu. Pro více porovnání bývá pro urychlení doporučována vhodná aproximace cyklů [20].



Obrázek 53 Dendrogram popisující vzdálenosti cyklů signálů.

Pro tento příklad byla jako prahová vzdálenost pro příslušnost do shluku zvolena hodnota 600. Ta se na základě mnoha porovnání jeví jako vyhovující i pro ostatní porovnání. Algoritmus dále na základě vzájemných vzdáleností ve shlucích vybere reprezentativní cyklus pro tento shluk. V tomto případě 4 rozdílné průběhy s různým počtem výskytů.



Obrázek 54 Průběhy vyskytující se v signálech.

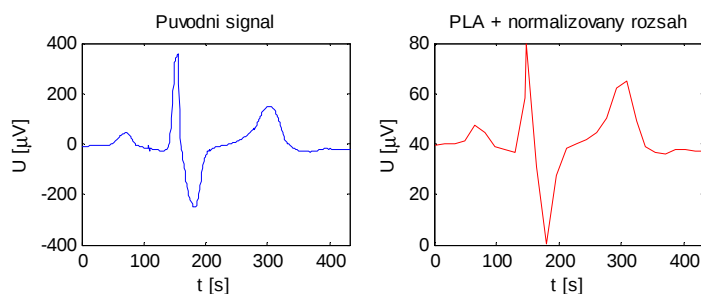
4.4 EKG segmentace

Již bylo uvedeno, že základem algoritmu automatizované segmentace EKG je sada vzorů se známými polohami významných klíčových bodů. Aby bylo možné vytvořit vzory a následně

hodnotit dosažené výsledky, omezíme se nadále pouze na referenční cykly, u kterých jsou klíčové polohy známe.

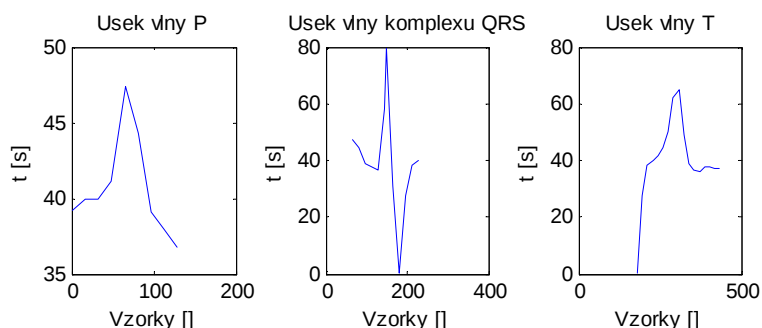
4.4.1 PLA

Po částech lineární aproximace je provedena funkcí *PLA.m* s krokem 32 ms a prahem 10 μV , jak bylo již uvedeno v teoretické části. Naneštěstí uzlové body klíčové body nekorrespondují přesně se zlomy aproximační linie, jak bylo předpokládáno v [8]. Jedná se o rozdíl do vzdálenosti 10 vzorků, většinou však kolem dvou až tří. A ani jiné nastavení výstup nezlepší.



Obrázek 55 Originalní signál a jeho normalizace a aproximace.

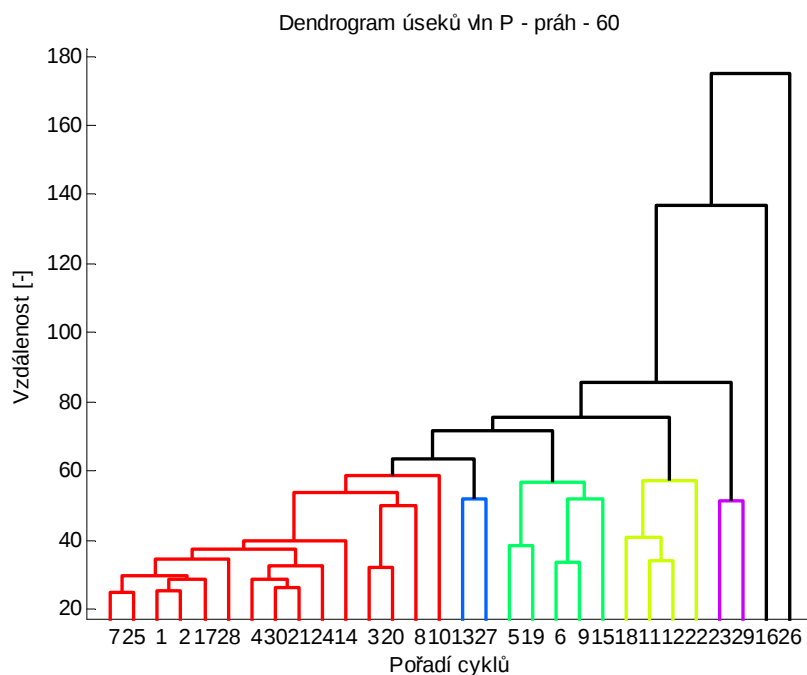
Dále je v tomto bloku zahrnuto požadované rozdělení na překrývající se segmenty odpovídající vlně P, komplexu QRS a vlně T.



Obrázek 56 Extrakce významných úseků z aproximovaného cyklu.

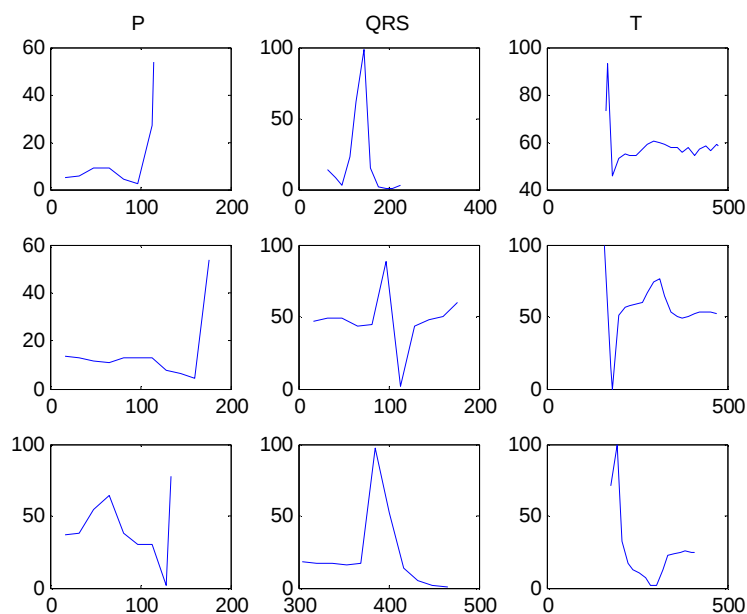
4.4.2 Vytvoření vzorů

Z dostupných signálů bude náhodně vybráno 60, z jejichž referenčních cyklů svodů I budou vytvořeny vzory. Jednotlivé aproximované úseky významných vln budou po určení vzájemných vzdáleností podrobeny shlukové analýze. Úseky reprezentující daný shluk budou uloženy. Pro úsek vlny P byla porovnáním výsledků jako hranice zvolena vzdálenost 60, pro komplex QRS 200 a 140 pro vlnu T.



Obrázek 57 Příklad dendrogramu při určování vzorů pro vlny P.

Pro vlnu P vzniklo 7 vzorů, pro komplex QRS 9 a vlnu T 5. Takováto „databáze“ by tedy teoreticky měla pokrýt $7 \times 9 \times 5 = 315$ různých průběhů cyklů EKG. Na vzorech bylo označeno pět klíčových bodů dostupných ze souboru MRESULTS (začátky a konce vlny P a komplexu QRS a konec vlny T). Výsledky byly uloženy do souborů PR.mat, QR.mat a TR.mat (Dodatek 3).



Obrázek 58 Příklad vytvořených vzorů pro segmentaci EKG.

4.4.3 Testování vytvořeného algoritmu

Algoritmus byl otestován pomocí skriptu *seg_test.m*. Segmentace byla provedena na množině cyklů, jež nebyly použity pro vytváření vzorů. Dosažené výsledky byly následně porovnány s očekávanými hodnotami. Výsledky jsou uloženy v *vysl_seg.mat* (Dodatek 3).

Tabulka 2 Vybrané výsledky testování algoritmu

	<i>začátek P</i>	<i>konec P</i>	<i>začátek QRS</i>	<i>konec QRS</i>	<i>konec T</i>
<i>Průměrný rozdíl</i>	14,09	6,31	7,56	7,15	14,28
<i>SD</i>	22,82	7,06	9,51	9,45	27,87
<i>Maximální rozdíl</i>	95	18	22	23	92

Zkoumáním výsledků testování algoritmu se projevuje nepříliš vysoká přesnost algoritmu. Zejména při určování začátků vln P a konců vln T. Procházením dílčích výsledků se potvrzuje, že vyšší odchylky jsou v oblasti konce P a QRS komplexu vzácnější, zatímco v oněch problematických místech jsou téměř pravidlem.

Závěr

Cílem diplomové práce bylo prostudování, pochopení principu a nabytí schopnosti vhodně využívat metody borcení časové osy v oblasti biologických signálů, poznat varianty algoritmu, navrhnout metody pro porovnání signálů EKG

Po úmyslně stručném úvodu do elektrokardiografie, následuje vlastní metoda DTW. Přes zdánlivou jednoduchost algoritmu lze dosáhnout zajímavých výsledků. V textu jsou popsány způsoby vytvoření matice vzdáleností, matice kumulativních vzdáleností i hledání výsledné cesty. Teoretická část je zakončena využitím DTW pro signály EKG. Z principu funkce již v počátcích vyplývala nutnost převážně porovnávat sekvence s předem vytvořenými vzory.

Úvod praktické části věnovaný předzpracování signálů se vesměs opírá o známé postupy bez širšího přibližování metod. V několika případech bylo dokonce využito převzatých programů. V této části se vyjma několika nedokonalých filtrací způsobených vysokým rušením prakticky nevyskytnul problém.

Algoritmus metody DTW samotné je poměrně jednoduchý. Při využití na kratších sekvencích v příkladech i dosti názorný. V rámci testování vyvstala otázka, jak určit vzdálenost, kterou již lze považovat za hraniční. Kapitola dále mimo jiné stanovuje omezující podmínky na tvar, průběh, způsob a oblast pohybu funkce, za účelem dosažení optimálního výsledku. V závěru této kapitoly je představena varianta borcení časové osy využívající odhadu první difference. Metoda zvaná též DDTW prokázala oproti klasické formě mnoho výhod, z nichž lze jmenovat například necitlivost na úroveň signálů.

Bylo prokázáno, že využitím globálních omezení je možné snížit výpočetní náročnost algoritmu. Při dalším zkoumání se výpočetní čas značně snížil i v případě aproximace signálu. Tento závěr však sám o sobě vyplývá z logiky věci.

Využití DTW jako metriky pro shlukovou analýzu prokázalo předpokládanou užitečnost zkoumané metody při prohlížení dlouhých úseků signálů, kdy ze záznamu obdržíme pouze reprezentativní cykly.

Metoda segmentace EKG bohužel uspokojivé výsledky nepřinesla. Již při aproximaci nedocházelo k naplnění předpokladů z literatury z hlediska očekávané polohy klíčových bodů ve zlomech aproximační linie. Tento počáteční, jak se ukázalo dále neodstranitelný, problém byl předzvěstí možných nepřesností. Zmíněný předpoklad se v průběhu implementace algoritmu potvrdil a metoda byla relativně nepřesná.

Již zmíněná nutnost neustálého porovnání s uloženými, předem vytvořenými vzory je jistým zklamáním i přesto, že využití DTW najde mnoho uplatnění. Stále to bude metrika pro jinou metodu, ale metrika užitečná.

Použité zdroje

- [1] KEOGH, E. J. a M. J. PAZZANI. *Derivative Dynamic Time Warping*. First SIAM International Conference on Data Mining, Chicago, 2001, s. 285–289.
- [2] PSUTKA, J.: *Komunikace s počítačem mluvenou řečí*. Vyd. 1. Praha: Academia, 1996. ISBN 80-200-0203-0.
- [3] VULLINGS, H. J. L. M., M. H. G. VERHAEGEN a H. B. VERBRUGGEN. *Automated ECG segmentation with Dynamic Time Warping*. In: *Proceedings of the 20th Annual International Conference of the IEEE, Hong Kong, 1998*. s. 163-166.
- [4] HUANG, B., KINSNER, W. ECG frame classification using dynamic time warping. Conference proceedings / Canadian Conference on Electrical and Computer Engineering, New Jersey, 2002. pp. 1105-1110.
- [5] SHEN, J., S., BAO, L.C., YANG a Y. LI. *The PLR-DTW Method for ECG Based Biometric Identification*. Annual International Conference of the IEEE, 2011, s. 5248-5251.
- [6] KEOGH, E. *Exact Indexing of Dynamic Time Warping*. University of California - Riverside Computer Science & Engineering Department. California. 2002.
- [7] KOZUMPLÍK, J. *Analýza biologických signálů*. Skripta FEKT VUT v Brně, 2011. 57 s.
- [8] ZIFAN, A., M. H. MORADI, S. SABERI a F. TOWHIDKHAH. *Automated Segmentation of ECG Signals using Piecewise Derivative Dynamic Time Warping*. Enformatika, Transactions on Engineering, Computing and Technology. 2006. s. 301-304.
- [9] VÍTEK, M. *Automatické rozměření signálu*. Brno, 2010. Disertační práce. FEKT VUT. Vedoucí dizertační práce: doc. Ing. Jiří Kozumplík CSc.
- [10] KLIMEK, M. aj. KOZUMPLÍK. *Shluková analýza elektrokardiografických signálů*. Elektrevue [online]. 2013, č. 4 [cit. 2013-12-27]. Dostupné z: <http://www.elektrevue.cz/cz/download/shlukova-analyza-elektrokardiograficky-signalu--ecg-cluster-analysis/>
- [11] TROJAN, S., SCHREIBER, M. *Atlas biologie člověka*. Praha: Scientia, 2007. 136 s.
- [12] KUBÁT, M., *Lineární a adaptivní filtry pro signály EKG*. Brno, 2012. Bakalářská práce. FEKT VUT. Vedoucí diplomové práce: doc. Ing. Jiří Kozumplík CSc.
- [13] ROMESBURG, H. *Cluster analysis for researchers*. Belmont, Calif.: Lifetime Learning Publications, c1984. ISBN 05-340-3248-6.
- [14] CUESTA FRAU, D. *Pattern Matching Techniques (Dynamic Time Warping)*. Valencia: Polytechnic University of Valencia, 2002
- [15] The CSE working party. *Recommendations for measurement standards in quantitative electrocardiography*. European Heart Journal, 1985, vol. 6, s. 815-825.

- [16] Mrázek, J. *Filtrace signálů EKG pomocí vlnkové transformace*. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií. Ústav biomedicínského a ekologického inženýrství, 2012. 55 s. Diplomová práce. Vedoucí práce: Ing. Lukáš Smital.
- [17] HAMPTON, John R. *EKG stručně, jasně, přehledně*. 2., rozš. vyd. Praha: Grada, 2005, 149 s. ISBN 80-247-0960-0.
- [18] FAYN, J., P. RUBEL. *Can the lessons learned from the assessment of automated electrocardiogram analysis in the Common Standards for quantitative Electrocardiography study benefit measurement of delayed contrast-enhanced magnetic resonance images?* Journal of Electrocardiography, 2007, vol. 40, s. 246-250.
- [19] HANSELMAN, Duane C. *Mastering MATLAB 6: a comprehensive tutorial and reference*. 1st Ed. Upper Saddle River, N.J.: Prentice-Hall, 2001, xviii, 814 s. ISBN 01-301-9468-9.
- [20] MÜLLER, Meinard. *Information retrieval for music and motion*. New York: Springer, 2007, xv, 313 p. ISBN 35-407-4047-3.
- [21] QIAN, Chen, G. Hu, F. Gu, P. Xiang. *Learning optimal warping window size of DTW for time series classification*. 11th International Conference on Information Science, Signal Processing and their Applications (ISSPA), 2012, s. 1272-1277.
- [22] KEOGH, E., RATANAMAHATANA Ch. A. *Everything you know about Dynamic Time Warping is Wrong*. Third Workshop on Mining Temporal and Sequential Data, International Conference on Knowledge Discovery and Data Mining, Seattle, 2004.

Seznam obrázků

Obrázek 1 Souvislost srdeční aktivity se signálem EKG. Podle [11], upraveno.....	9
Obrázek 2 Náčrt spekter základních vln a kmitů EKG.....	9
Obrázek 3 Příklad rozměření signálu EKG [10].....	11
Obrázek 4 Využití Euklidovské vzdálenosti a DTW. Převzato z[6], upraveno.....	12
Obrázek 5 Příklad nesplnění okrajových podmínek [14].	14
Obrázek 6 Následek nemonotónní výsledné cesty [14].....	14
Obrázek 7 Následek nespojitosti výsledné cesty [14].....	14
Obrázek 8 Příklad bortivé cesty [1].	15
Obrázek 9 Následek mnoha kroků ve vertikálním nebo horizontálním směru [14].....	16
Obrázek 10 Možné směry pohybu funkce (a) pro rovnici 11, (b) pro 13 a (c) pro 15 [19].....	16
Obrázek 11 Tři bortivé cesty s ohledem na rozdílný krok [19].	17
Obrázek 12 Následek pohybu funkce daleko od diagonály[14].	17
Obrázek 13 Příklad omezujících podmínek. Podle [2].	18
Obrázek 14 Zarovnání dvou identických posloupností [1].....	20
Obrázek 15 Vznik singularit při mírné změně [1]	20
Obrázek 16 Příklad klasifikovaného signálu [10].....	23
Obrázek 17 Dendrogram příslušný klasifikovanému signálu [10].....	24
Obrázek 18 PLA: (a) aproximace v pořádku, (b) bod je za hranicí, úsečka musí být zkrácena [3].	24
Obrázek 19 Příklad rozložení jednoho aproximovaného cyklu EKG [3].	25
Obrázek 20 Příklady referenčních vln, jejichž kombinací lze vytvořit 27 různých průběhů [3]..	25
Obrázek 21 Svod I záznamu MO1_001_12 a odstranění rušení.....	28
Obrázek 22 Detekce QRS	28
Obrázek 23 Kontrola rozměrovacího algoritmu	29
Obrázek 24 Úseky porovnávaných signálů.....	29
Obrázek 25 Matice lokálních vzdáleností.....	30
Obrázek 26 Matice kumulovaných vzdáleností s vyznačenou cestou.	30
Obrázek 27 Signály po borcení časové osy	31
Obrázek 28 DTW s čísly.....	31
Obrázek 29 DTW pro dva cykly EKG.....	31

Obrázek 30 Výsledná cesta při zarovnávání vybraných cyklů EKG.....	32
Obrázek 31 DTW při různém rozměření téhož cyklu.....	33
Obrázek 32 Nesprávné zarovnání signálů.....	33
Obrázek 33 Bortivá cesta pro referenční cyklus signálu MO1_115_12.....	34
Obrázek 34 Změna v signálu.	34
Obrázek 35 Zarovnané signály,	35
Obrázek 36 Bortivá cesta.	35
Obrázek 37 Zarovnání signálu MO1_115_12 s krokem podle rovnice 13.....	36
Obrázek 38 Bortivá cesta při využití kroku (b).	36
Obrázek 39 Zarovnání při kroku (c).	36
Obrázek 40 Bortivá cesta při kroku (c).....	37
Obrázek 41 Zarovnání signálů.	38
Obrázek 42 Výsledná cesta.....	38
Obrázek 43 Reakce algoritmu na změnu při využití kroku (c).....	38
Obrázek 44 Signály s největším rozdílem vzdáleností s ohledem na použitou šířku okna	39
Obrázek 45 Bortivá cesta pro globálním omezení s šířkou okna na 20% šířky diagonály.....	40
Obrázek 46 Bortibá cesta při globálním omezení s šířkou okna na 20% šířky diagonály.....	40
Obrázek 47 Signály se zborcenou časovou osou při globálním šířkou okna na 10% šířky diagonály.....	40
Obrázek 48 Signály se zborcenou časovou osou při globálním šířkou okna na 20% šířky diagonály.....	41
Obrázek 49 Signál a odhad jeho první difference podle 28.	41
Obrázek 50 DDTW signálu MO1_115_12.....	42
Obrázek 51 Bortivá cesta při zarovnání signálu MO1_115_12 s různými úrovněmi.....	42
Obrázek 52 Signály vybrané pro shlukovou analýzu	43
Obrázek 53 Dendrogram popisující vzdálenosti cyklů signálů.	44
Obrázek 54 Průběhy vyskytující se v signálech.	44
Obrázek 55 Originální signál a jeho normalizace a aproximace.	45
Obrázek 56 Extrakce významných úseků z aproximovaného cyklu.	45
Obrázek 57 Příklad dendrogramu při určování vzorů pro vlny P.....	46
Obrázek 58 Příklady vytvořených vzorů pro segmentaci EKG.....	46

Seznam zkratek

EKG	Elektrokardiogram
DTW	Metoda borcení časové osy
t_R	Pozice R kmitu
A, B	Posloupnosti vzorků
a_i	Prvek posloupnosti
$\psi(i)$	Cesta maticí vzdáleností
w_K	Jiné označení cesty (prvku cesty)
$\delta(i, j)$	Prvek matice vzdáleností
$\gamma(i, j)$	Prvek matice kumulativních vzdáleností
ω	Šířka okna
$e_{(ab)c}$	Vzdálenost shluku (ab) od elementu c
$t(a_i)$	Doba trvání aproximační linky

Obsah přiloženého CD

Název	Popis
Milan_Kubát_DP.pdf	Vlastní text práce
Zdrojové kódy a signály	Navržené funkce, použité filtrované signály
Obrázky	Grafické výstupy
Hodnocení	.mat soubory výsledků

Dodatky

Dodatek 1 Seznam vyřazených signálů

MO1_006_12	MO1_047_12	MO1_089_12	MO1_109_12
MO1_010_12	MO1_050_12	MO1_092_12	MO1_111_12
MO1_012_12	MO1_052_12	MO1_093_12	MO1_112_12
MO1_018_12	MO1_057_12	MO1_094_12	MO1_120_12
MO1_032_12	MO1_065_12	MO1_095_12	
MO1_034_12	MO1_067_12	MO1_100_12	
MO1_045_12	MO1_070_12	MO1_105_12	

Dodatek 2 Kódy

Funkce nulovani

```
[filtrovany_signal]=nulovani(x)
```

Jediným vstupem funkce je signál postižený šumovou složkou. Nejdůležitější částí je převedení do spektrální oblasti a vynulování příslušných spektrálních čar.

```
X = fft(x); % spektrum vstupního signalu
df = fvz/N; % vzdálenost spektrálních
čar
hp = ceil(fhp/df); % vypočet
od = ceil(fdmez/df); % mezí
do = ceil(fhmez/df); % na frekvenční ose
Y = X;

Y(1:hp+1) = 0; % vynulování příslušných
spektrálních čar HP
Y(N-hp+1:end) = 0; % vynulování symetrických čar
Y(od+1:do+1) = 0; % vynulování příslušných
spektrálních čar PZ
Y(N-do+1:N-od+1) = 0; % vynulování symetrických čar
filtrovany_signal = real(ifft(Y));
```

Funkce dtw

```
[r s vzd]=dtw(A, B, krok, fvz, DDTW, vykreslení, GL)
```

Vstupem jsou úseky k porovnání A a B, krok při výpočtu matice kumulativních vzdáleností, vzorkovací frekvence, požadavek na DDTW, vykreslení a šířka globálního omezení v procentech.

```
Mvz=abs(repmat(A1',1,n)-repmat(B1,m,1));
```

Výpočet matice vzdáleností probíhá za pomoc funkce *repmat*, která „roztáhne“ vektory do matice. Matice jsou pak odečteny. Matice kumulovaných vzdáleností je počítána ve dvou for cyklech.

```
for i=2:m
    for j=2:n
        MAT(i,j)=Mvz(i,j)+min([MAT(i-1,j-1), MAT(i-1,j), MAT(i,j-1)]);
    end
end
```

Je-li nastaveno globální omezení, pohybujeme se pomocí *while* po úhlopříčce, za současného zohlednění rozdílu délek signálů (*posun*).

```
while i+j<m+n+2
    MAT(i,j)=Mvz(i-1,j-1)+min([MAT(i-1,j-1), MAT(i,j-1), MAT(i-1,j)]);
% hodnota na uhlopricce
    for t=j+1:min(j+G1-1, n+1) % vyplni radky
        MAT(i,t)=Mvz(i-1,t-1)+min([MAT(i-1,t-1), MAT(i,t-1), MAT(i-1,
t)]);
    end
    for r=i+1:min(i+G2-1,m+1) % sloupce
        MAT(r,j)=Mvz(r-1,j-1)+min([MAT(r-1,j-1), MAT(r,j-1), MAT(r-1,
j)]);
    end

    if sum(posun1(2,1)==j)>0 % uprava indexu s ohledem na prubeh
uhlopricky
        j=j-1;
        posun1(:,1)=[];
    elseif sum(posun1(1,1)==i)>0
        i=i-1;
        posun1(:,1)=[];
    end
    if i~=m+1; i=i+1; end % iterace
    if j~=n+1; j=j+1; end
end
```

Zpětné trasování cesty už je poté jen ukládání indexů do proměnných *r* a *s*, při pohybu pi prvcích s nejmenší kumulovanou vzdáleností.

Funkce *shlu_anal.m*

```
[Ref] = shlu_anal(signaly, svody, metoda, krok, prah, vykresleni)
```

Funkce shlukuje cykly vybraných signálů do shluků podle zvoleného prahu. Do proměnných *signaly* a *svody* se zadávají pouze čísla signálů (svodů).

Funkce PLA.m

[aprs, P, Q, T]=PLA(signal, n, th, fvz, vykresleni, QRS)

Provádí po částech lineární aproximaci zadaného cyklu. A rozdělení na úseky vlny P, komplexu QRS a vlny T.

```
while ii<=N+1
    if i+ms<=N; ii=i+ms; else ii=N; end % krok
    ok=0; % flag
    while ok~=1
        k=(signal(ii)-signal(i))/(ii-i); % směrnice přímky spojující
    body
        q1=signal(i)-k*i; % posunutí 1
        for im=i+1:ii % kontrola nepřekročení
prahu
            for im=ii:-1:i+1
                q2=signal(im)-k*im; % posunutí 2 - rovnoběžné
přímky procházející bodem
                vzd=abs(q1-q2)*sin((pi/2)-abs(atan(k))); % výpočet
vzdálenosti
                if vzd > th % překročení prahu?
                    ii=im; % ANO - propojit s tímto
bodem
                    break % a ukončit kontrolu prahu
                elseif im==ii % Prohledané celé - další
úsek
                    ok=1; % flag
                    i=ii; % začátek nového úseku
                    aprs(1:4,end+1)=[signal(ii);ii; k; ii-aprs(2,end)]; %
uložení výstupu
                end
            end
        end
        if ii==N; break; end
    end
```

Skript seg_test.m

Po načtení testovací množiny a vzorů, se vzory porovná pomocí funkce dtw_seg jednotlivé cykly. Následně určí nejmenší vzdálenost a polohy klíčových bodů.

Funkce dtw_seg

[vzd, r, s]=dtw_seg(A, B)

Je prakticky shodná s dtw. S ohledem na fakt, že využívá speciální krok byla pro přehlednost vyčleněna zvlášť.

Dodatek 3 Popisy uložených souborů

delka.mat

V řádcích jsou hodnoty uloženy v pořadí číslo signálu, typ změny délky, vzdálenost, střední hodnota nezměněného cyklu, jeho délka, střední hodnota změněného a jeho délka a maximální počet kroků v horizontálním a vertikálním směru.

uroven.mat, zmena.mat, uroven_k2.mat, zmena_k2.mat,
zmena_k3.mat

Název před podtržítkem udává, jde-li o změnu u úrovní signálu nebo o vložení změny do něj. Číslovka u k pak vyjadřuje použitý krok.

V řádcích jsou hodnoty uloženy v pořadí číslo signálu, typ změny délky, vzdálenost, střední hodnota nezměněného cyklu, jeho délka, střední hodnota změněného a jeho délka a maximální počet kroků v horizontálním a vertikálním směru.

CPU.mat

V řádcích jsou data uložena v pořadí název signálu, čas výpočtu s omezením, bez omezení a procentuální rozdíl časů.

vysl_seg.mat

Výsledky jsou uloženy v buňkovém poli. První sloupeček obsahuje čísla signálů, druhý samotné výsledky. Ty jsou v dvouřádkové matici, kdy první řádek obsahuje očekávané hodnoty klíčových bodů a druhý hodnoty určené v tomto pořadí po sloupcích: začátek a konec vlny P, začátek a konec komplexu QRS, konec vlny T.